GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: January 15, 2003, 20:35:40 ; Search time 32.5 Seconds
(without alignments)
1371.902 Million cell updates/sec

Title: L00394_COPY_1_100
Perfect score: 100
Sequence: 1 CTCCTTTGGCAGTCACACGG.....GTGGTGTGCTCCTGCGCCCG 100

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 393868 seqs, 222934149 residues

Total number of hits satisfying chosen parameters: 787736

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	54	54.0	216	10	US-09-960-352-12506	Sequence 12506, A
	2	54	54.0	424	10	US-09-960-352-5863	Sequence 5863, Ap
c	3	30.8	30.8	295	10	US-09-867-701-3188	Sequence 3188, Ap
	4	29.6	29.6	7132	12	US-10-044-090-780	Sequence 780, App
c	5	28.2	28.2	707	9	US-10-001-189-40	Sequence 40, Appl
c	6	28.2	28.2	3662	9	US-10-001-189-41	Sequence 41, Appl
c	7	28.2	28.2	4613	9	US-10-001-189-46	Sequence 46, Appl
	8	28.2	28.2	4941	9	US-10-001-189-53	Sequence 53, Appl
c	9	28.2	28.2	4943	9	US-10-001-189-54	Sequence 54, Appl
c	10	28.2	28.2	4944	9	US-10-001-189-55	Sequence 55, Appl
c	11	28.2	28.2	4944	9	US-10-001-189-56	Sequence 56, Appl
	12	28.2	28.2	4951	9	US-10-001-189-51	Sequence 51, Appl
	13	28.2	28.2	4952	9	US-10-001-189-52	Sequence 52, Appl
	14	28.2	28.2	8999	9	US-10-001-189-48	Sequence 48, Appl
	15	28.2	28.2	9012	9	US-10-001-189-49	Sequence 49, Appl
	16	28.2	28.2	9013	9	US-10-001-189-50	Sequence 50, Appl
c	17	27	27.0	417	10	US-09-864-761-3885	Sequence 3885, Ap
c	18	27	27.0	666	9	US-09-738-626-706	Sequence 706, App
	19	27	27.0	14849	10	US-09-873-403-1	Sequence 1, Appli
	20	26.8	26.8	3177	10	US-09-898-570-15	Sequence 15, Appl
	21	26.8	26.8	4550	10	US-09-898-570-13	Sequence 13, Appl
c	22	26.8	26.8	4821	10	US-09-898-570-11	Sequence 11, Appl
	23	26.6	26.6	783	9	US-10-174-590-341	Sequence 341, App
	24	26.6	26.6	783	9	US-10-176-758-341	Sequence 341, App
	25	26.6	26.6	783	12	US-10-052-586-341	Sequence 341, App
	26	26.6	26.6	2002	10	US-09-935-390A-5	Sequence 5, Appli
	27	26.4	26.4	4440	10	US-09-898-570-9	Sequence 9, Appli
c	28	26	26.0	1491	10	US-09-796-393-1	Sequence 1, Appli
c	29	25.6	25.6	449	10	US-09-867-701-6811	Sequence 6811, Ap
	30	25.6	25.6	2264	9	US-10-004-717-49	Sequence 49, Appl
	31	25.4	25.4	286	10	US-09-934-586A-21	Sequence 21, Appl
	32	25.4	25.4	1381	9	US-09-853-526-72	Sequence 72, Appl
	33	25.4	25.4	1381	9	US-09-853-526-184	Sequence 184, App
	34	25.4	25.4	1381	10	US-09-901-484A-72	Sequence 72, Appl
	35	25.4	25.4	1381	10	US-09-901-484A-184	Sequence 184, App
	36	25.4	25.4	4550	9	US-09-853-526-182	Sequence 182, App
	37	25.4	25.4	4550	10	US-09-901-484A-182	Sequence 182, App
	38	25.2	25.2	233	10	US-09-864-761-24005	Sequence 24005, A
c	39	25.2	25.2	405	9	US-10-040-739-547	Sequence 547, App
	40	25.2	25.2	589	10	US-09-864-761-7274	Sequence 7274, Ap
	41	25.2	25.2	1668	9	US-09-738-626-2694	Sequence 2694, Ap
	42	25	25.0	488	10	US-09-833-381-1066	Sequence 1066, Ap
c	43	25	25.0	576	9	US-09-764-868-523	Sequence 523, App
c	44	25	25.0	1461	10	US-09-759-990-1	Sequence 1, Appli
	45	25	25.0	1837	10	US-09-925-301-365	Sequence 365, App

ALIGNMENTS

RESULT 1
US-09-960-352-12506
; Sequence 12506, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AI
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 12506
; LENGTH: 216
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 53-LIB34-065-Q1-E1-F2
US-09-960-352-12506

Query Match 54.0%; Score 54; DB 10; Length 216;
Best Local Similarity 76.7%; Pred. No. 5e-09;
Matches 66; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

Qy 14 CACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGAACA 73
||| || || |||||||||||||||| |||| ||| |||||||||||| ||||| |
Db 58 CACGCGTGAGATCTGCAGCCTGGACAATGGGGGCTGCGACCAGTTCTGCAGGGAGGAGCG 117

Qy 74 GAACTCTGTGGTGTGCTCCTGCGCCC 99
| | ||| |||||||||||| |
Db 118 CAGCGAGGTGCGGTGCTCCTGCGCGC 143

RESULT 2
US-09-960-352-5863
; Sequence 5863, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing

```
;  APPLICANT:  Byatt, John C.
;  APPLICANT:  Mathialagan, Nagappan
;  TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
;  TITLE OF INVENTION:  MUSCLE AND FAT DEPOSITION
;  FILE REFERENCE: 16511.006/37-21(10298)C
;  CURRENT APPLICATION NUMBER: US/09/960,352
;  CURRENT FILING DATE:  2001-09-24
;  NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 5863
;   LENGTH: 424
;   TYPE: DNA
;   ORGANISM: Bos taurus
;   OTHER INFORMATION: Clone ID: 25-LIB34-026-Q1-E1-G1
US-09-960-352-5863

  Query Match             54.0%;  Score 54;  DB 10;  Length 424;
  Best Local Similarity   76.7%;  Pred. No. 6e-09;
  Matches   66;  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

Qy      14 CACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGAACA 73
           ||| ||  || ||||||||||||||| |||| ||| ||||||||||||   ||||| |
Db      11 CACGCGTGAGATCTGCAGCCTGGACAATGGGGGCTGCGACCAGTTCTGCAGGGAGGAGCG 70

Qy      74 GAACTCTGTGGTGTGCTCCTGCGCCC 99
             | |   |||  |||||||||||| |
Db      71 CAGCGAGGTGCGGTGCTCCTGCGCGC 96


RESULT 3
US-09-867-701-3188/c
; Sequence 3188, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
;  APPLICANT: Aglate, Paul A.
;  APPLICANT:  Jones, Robert
;  APPLICANT:  Harlocker, Susan L.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF OVARIAN CANCER
;  FILE REFERENCE: 210121.497
;  CURRENT APPLICATION NUMBER: US/09/867,701
;  CURRENT FILING DATE:  2001-05-29
;  NUMBER OF SEQ ID NOS: 10912
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3188
;   LENGTH: 295
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-867-701-3188

  Query Match             30.8%;  Score 30.8;  DB 10;  Length 295;
  Best Local Similarity   63.5%;  Pred. No. 0.24;
  Matches   47;  Conservative    0;  Mismatches   27;  Indels    0;  Gaps    0;

Qy      27 TGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGAACAGAACTCTGTGGTG 86
           ||||||| ||| ||||||| |||   |||| |||||||| |  | | |       |
Db     108 TGCAGCCAGGAGAACGGGGGCTGCCTCCAGATCTGCCACAACAAGCCGGGTAGCTTCCAC 49

Qy      87 TGCTCCTGCGCCCG 100
           || ||||||  | |
Db      48 TGTTCCTGCCACAG 35


RESULT 4
US-10-044-090-780
; Sequence 780, Application US/10044090
; Patent No. US20020137081A1
; GENERAL INFORMATION:
;  APPLICANT: Olga Bandman
;  TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
;  FILE REFERENCE: PA-0028 US
;  CURRENT APPLICATION NUMBER: US/10/044,090
;  CURRENT FILING DATE:  2002-01-09
;  NUMBER OF SEQ ID NOS: 850
;  SOFTWARE: PERL Program
; SEQ ID NO 780
;   LENGTH: 7132
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20020137081A1 198499.13
US-10-044-090-780

  Query Match             29.6%;  Score 29.6;  DB 12;  Length 7132;
  Best Local Similarity   61.8%;  Pred. No. 1.4;
  Matches   47;  Conservative    0;  Mismatches   29;  Indels    0;  Gaps    0;

Qy      20 GAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGAACAGAACTC 79
           | ||  |||||   |  |||||||| | ||||| ||||   |||||   ||  | |    |
Db      78 GGAGAACTGCAATGTTAACAACGGTGGCTGTGCCCAGAAGTGCCAGATGGTGCGGGGGGC 137

Qy      80 TGTGGTGTGCTCCTGC 95
             |||  |||  |||||
Db     138 AGTGCAGTGTACCTGC 153


RESULT 5
US-10-001-189-40/c
; Sequence 40, Application US/10001189
; Patent No. US20020173634A1
; GENERAL INFORMATION:
;  APPLICANT: FRASER JR., MALCOLM J.
;  APPLICANT:  LI, XU
;  APPLICANT:  BEAM, TERESA
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
;  TITLE OF INVENTION:  MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
;  TITLE OF INVENTION:  VECTOR PIGGYBAC
;  FILE REFERENCE: 835910-92098
;  CURRENT APPLICATION NUMBER: US/10/001,189
;  CURRENT FILING DATE:  2001-10-30
;  PRIOR APPLICATION NUMBER: 60/244,984
;  PRIOR FILING DATE: 2000-11-01
;  PRIOR APPLICATION NUMBER: 60/244,677
;  PRIOR FILING DATE: 2000-10-31
;  NUMBER OF SEQ ID NOS: 70
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 40
;   LENGTH: 707
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: ITR Cartridge
;   OTHER INFORMATION: sequence
US-10-001-189-40

  Query Match             28.2%;  Score 28.2;  DB 9;  Length 707;
  Best Local Similarity   61.6%;  Pred. No. 2.2;
  Matches   45;  Conservative    0;  Mismatches   28;  Indels    0;  Gaps    0;

Qy      18 CGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGAACAGAAC 77
           | || || || | | |||  || | || |||||| || |    ||| ||| |||| |
Db     358 CTGATGCGCTTCGGGCTGACCATCCGGAACTGTGTCCGGAAAAGCCGCGACGAACTGGTA 299

Qy      78 TCTGTGGTGTGCT 90
           ||   ||||  ||
Db     298 TCCCAGGTGGCCT 286


RESULT 6
US-10-001-189-41/c
; Sequence 41, Application US/10001189
; Patent No. US20020173634A1
; GENERAL INFORMATION:
;  APPLICANT: FRASER JR., MALCOLM J.
;  APPLICANT:  LI, XU
;  APPLICANT:  BEAM, TERESA
```

```
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
;  TITLE OF INVENTION:  MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
;  TITLE OF INVENTION:  VECTOR PIGGYBAC
;  FILE REFERENCE: 835910-92098
;  CURRENT APPLICATION NUMBER: US/10/001,189
;  CURRENT FILING DATE:  2001-10-30
;  PRIOR APPLICATION NUMBER: 60/244,984
;  PRIOR FILING DATE: 2000-11-01
;  PRIOR APPLICATION NUMBER: 60/244,677
;  PRIOR FILING DATE: 2000-10-31
;  NUMBER OF SEQ ID NOS: 70
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
;   LENGTH: 3662
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: pXL-Bac
;   OTHER INFORMATION: sequence
S-10-001-189-41

  Query Match             28.2%;  Score 28.2;  DB ;  Length 3662;
  Best Local Similarity   61.6%;  Pred. No. 3.5;
  Matches   45;  Conservative    0;  Mismatches   28;  Indels    0;  Gaps    0;

Qy      18 CGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGAACAGAAC 77
           | || || || | | |||  || | || |||||| || |    ||| ||| |||| |
Db    1339 CTGATGCGCTTCGGGCTGACCATCCGGAACTGTGTCCGGAAAAGCCGCGACGAACTGGTA 1280

Qy      78 TCTGTGGTGTGCT 90
           ||   ||||  ||
Db    1279 TCCCAGGTGGCCT 1267


RESULT 7
US-10-001-189-46/c
; Sequence 46, Application US/10001189
; Patent No. US20020173634A1
; GENERAL INFORMATION:
;  APPLICANT: FRASER JR., MALCOLM J.
;  APPLICANT:  LI, XU
;  APPLICANT:  BEAM, TERESA
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
;  TITLE OF INVENTION:  MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
;  TITLE OF INVENTION:  VECTOR PIGGYBAC
;  FILE REFERENCE: 835910-92098
;  CURRENT APPLICATION NUMBER: US/10/001,189
;  CURRENT FILING DATE:  2001-10-30
   PRIOR APPLICATION NUMBER: 60/244,984
;  PRIOR FILING DATE: 2000-11-01
;  PRIOR APPLICATION NUMBER: 60/244,677
;  PRIOR FILING DATE: 2000-10-31
;  NUMBER OF SEQ ID NOS: 70
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 46
;   LENGTH: 4613
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: pCRII-ITR
;   OTHER INFORMATION: sequence
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (344)..(922)
US-10-001-189-46

  Query Match             28.2%;  Score 28.2;  DB 9;  Length 4613;
  Best Local Similarity   61.6%;  Pred. No. 3.7;
  Matches   45;  Conservative    0;  Mismatches   28;  Indels    0;  Gaps    0;

Qy      18 CGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGAACAGAAC 77
           | || || || | | |||  || | || |||||| || |    ||| ||| |||| |
Db     651 CTGATGCGCTTCGGGCTGACCATCCGGAACTGTGTCCGGAAAAGCCGCGACGAACTGGTA 592

Qy      78 TCTGTGGTGTGCT 90
           ||   ||||  ||
Db     591 TCCCAGGTGGCCT 579


RESULT 8
US-10-001-189-53
; Sequence 53, Application US/10001189
; Patent No. US20020173634A1
; GENERAL INFORMATION:
;  APPLICANT: FRASER JR., MALCOLM J.
;  APPLICANT:  LI, XU
;  APPLICANT:  BEAM, TERESA
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
;  TITLE OF INVENTION:  MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
;  TITLE OF INVENTION:  VECTOR PIGGYBAC
;  FILE REFERENCE: 835910-92098
;  CURRENT APPLICATION NUMBER: US/10/001,189
;  CURRENT FILING DATE:  2001-10-30
;  PRIOR APPLICATION NUMBER: 60/244,984
;  PRIOR FILING DATE: 2000-11-01
;  PRIOR APPLICATION NUMBER: 60/244,677
;  PRIOR FILING DATE: 2000-10-31
;  NUMBER OF SEQ ID NOS: 70
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 53
;   LENGTH: 4941
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: pXL-Bac-ECFP
;   OTHER INFORMATION: sequence
US-10-001-189-53

  Query Match             28.2%;  Score 28.2;  DB 9;  Length 4941;
  Best Local Similarity   61.6%;  Pred. No. 3.8;
  Matches   45;  Conservative    0;  Mismatches   28;  Indels    0;  Gaps    0;

Qy      18 CGGAAGCTCTGCAGCCTGG..CAACGGGGACTGTGACCAGTTCTGCCACGAGGAACAGAAC 77
           | || || || | | |||  || | || |||||| || |    ||| ||| |||| |
Db     778 CTGATGCGCTTCGGGCTGACCATCCGGAACTGTGTCCGGAAAAGCCGCGACGAACTGGTA 837

Qy      78 TCTGTGGTGTGCT 90
           ||   ||||  ||
Db     838 TCCCAGGTGGCCT 850


RESULT 9
US-10-001-189-54/c
; Sequence 54, Application US/10001189
; Patent No. US20020173634A1
; GENERAL INFORMATION:
;  APPLICANT: FRASER JR., MALCOLM J.
;  APPLICANT:  LI, XU
;  APPLICANT:  BEAM, TERESA
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
;  TITLE OF INVENTION:  MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
;  TITLE OF INVENTION:  VECTOR PIGGYBAC
;  FILE REFERENCE: 835910-92098
;  CURRENT APPLICATION NUMBER: US/10/001,189
;  CURRENT FILING DATE:  2001-10-30
;  PRIOR APPLICATION NUMBER: 60/244,984
;  PRIOR FILING DATE: 2000-11-01
;  PRIOR APPLICATION NUMBER: 60/244,677
;  PRIOR FILING DATE: 2000-10-31
;  NUMBER OF SEQ ID NOS: 70
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 54
;   LENGTH: 4943
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
```

```
;   OTHER INFORMATION: Description of Artificial Sequence: PBS-ITR-ECFP
;   OTHER INFORMATION: sequence
US-10-001-189-54

  Query Match             28.2%;  Score 28.2;  DB 9;  Length 4943;
  Best Local Similarity   61.6%;  Pred. No. 3.8;
  Matches   45;  Conservative    0;  Mismatches   28;  Indels    0;  Gaps    0;

Qy      18 CGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGAACAGAAC 77
           | || || || | | |||  || | || |||||| || |    ||| ||| |||| |
Db    1076 CTGATGCGCTTCGGGCTGACCATCCGGAACTGTGTCCGGAAAAGCCGCGACGAACTGGTA 1017

Qy      78 TCTGTGGTGTGCT 90
           ||   ||||  ||
Db    1016 TCCCAGGTGGCCT 1004


RESULT 10
US-10-001-189-55/c
; Sequence 55, Application US/10001189
  Patent No. US20020173634A1
  GENERAL INFORMATION:
;  APPLICANT: FRASER JR., MALCOLM J.
;  APPLICANT:  LI, XU
;  APPLICANT:  BEAM, TERESA
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
;  TITLE OF INVENTION:  MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
;  TITLE OF INVENTION:  VECTOR PIGGYBAC
;  FILE REFERENCE: 835910-92098
;  CURRENT APPLICATION NUMBER: US/10/001,189
;  CURRENT FILING DATE:  2001-10-30
;  PRIOR APPLICATION NUMBER: 60/244,984
;  PRIOR FILING DATE: 2000-11-01
;  PRIOR APPLICATION NUMBER: 60/244,677
;  PRIOR FILING DATE: 2000-10-31
;  NUMBER OF SEQ ID NOS: 70
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 55
;   LENGTH: 4944
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: PBS-ITR-EGFP
;   OTHER INFORMATION: sequence
US-10-001-189-55

  Query Match             28.2%;  Score 28.2;  DB 9;  Length 4944;
  Best Local Similarity   61.6%;  Pred. No. 3.8;
  Matches   45;  Conservative    0;  Mismatches   28;  Indels    0;  Gaps    0;

Qy      18 CGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGAACAGAAC 77
           | || || || | | |||  || | || |||||| || |    ||| ||| |||| |
Db    1076 CTGATGCGCTTCGGGCTGACCATCCGGAACTGTGTCCGGAAAAGCCGCGACGAACTGGTA 1017

Qy      78 TCTGTGGTGTGCT 90
           ||   ||||  ||
Db    1016 TCCCAGGTGGCCT 1004


RESULT 11
US-10-001-189-56/c
; Sequence 56, Application US/10001189
; Patent No. US20020173634A1
; GENERAL INFORMATION:
;  APPLICANT: FRASER JR., MALCOLM J.
;  APPLICANT:  LI, XU
;  APPLICANT:  BEAM, TERESA
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
;  TITLE OF INVENTION:  MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
;  TITLE OF INVENTION:  VECTOR PIGGYBAC
;  FILE REFERENCE: 835910-92098
;  CURRENT APPLICATION NUMBER: US/10/001,189
;  CURRENT FILING DATE:  2001-10-30
;  PRIOR APPLICATION NUMBER: 60/244,984
;  PRIOR FILING DATE: 2000-11-01
;  PRIOR APPLICATION NUMBER: 60/244,677
;  PRIOR FILING DATE: 2000-10-31
;  NUMBER OF SEQ ID NOS: 70
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 56
;   LENGTH: 4944
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: pBS-ITR-EYFP
;   OTHER INFORMATION: sequence
US-10-001-189-56

  Query Match             28. %;  Score 28.2;  DB 9;  Length 4944;
  Best Local Similarity   61.6%;  Pred. No. 3.8;
  Matches   45;  Conservative    0;  Mismatches   28;  Indels    0;  Gaps    0;

Qy      18 CGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGAACAGAAC 77
           | || || || | | |||  || | || |||||| || |    ||| ||| |||| |
Db    1076 CTGATGCGCTTCGGGCTGACCATCCGGAACTGTGTCCGGAAAAGCCGCGACGAACTGGTA 1017

Qy      78 TCTGTGGTGTGCT 90
           ||   ||||  ||
Db    1016 TCCCAGGTGGCCT 1004


RESULT 12
US-10-001-189-51
; Sequence 51, Application US/10001189
; Patent No. US20020173634A1
; GENERAL INFORMATION:
;  APPLICANT: FRASER JR., MALCOLM J.
;  APPLICANT:  LI, XU
;  APPLICANT:  BEAM, TERESA
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
;  TITLE OF INVENTION:  MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
;  TITLE OF INVENTION:  VECTOR PIGGYBAC
;  FILE REFERENCE: 835910-92098
;  CURRENT APPLICATION NUMBER: US/10/001,189
;  CURRENT FILING DATE:  2001-10-30
;  PRIOR APPLICATION NUMBER: 60/244,984
;  PRIOR FILING DATE: 2000-11-01
;  PRIOR APPLICATION NUMBER: 60/244,677
;  PRIOR FILING DATE: 2000-10-31
;  NUMBER OF SEQ ID NOS: 70
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 51
;   LENGTH: 4951
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: pXL-Bac-EYFP
;   OTHER INFORMATION: sequence
US-10-001-189-51

  Query Match             28.2%;  Score 28.2;  DB 9;  Length 4951;
  Best Local Similarity   61.6%;  Pred. No. 3.8;
  Matches   45;  Conservative    0;  Mismatches   28;  Indels    0;  Gaps    0;

Qy      18 CGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGAACAGAAC 77
           | || || || | | |||  || | || |||||| || |    ||| ||| |||| |
Db     778 CTGATGCGCTTCGGGCTGACCATCCGGAACTGTGTCCGGAAAAGCCGCGACGAACTGGTA 837

Qy      78 TCTGTGGTGTGCT 90
           ||   ||||  ||
Db     838 TCCCAGGTGGCCT 850


RESULT 13
US-10-001-189-52
; Sequence 52, Application US/10001189
```

```
; Patent No. US20020173634A1
; GENERAL INFORMATION:
;  APPLICANT: FRASER JR., MALCOLM J.
;  APPLICANT:  LI, XU
;  APPLICANT:  BEAM, TERESA
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
;  TITLE OF INVENTION:  MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
;  TITLE OF INVENTION:  VECTOR PIGGYBAC
;  FILE REFERENCE: 835910-92098
;  CURRENT APPLICATION NUMBER: US/10/001,189
;  CURRENT FILING DATE:  2001-10-30
;  PRIOR APPLICATION NUMBER: 60/244,984
;  PRIOR FILING DATE: 2000-11-01
;  PRIOR APPLICATION NUMBER: 60/244,677
;  PRIOR FILING DATE: 2000-10-31
;  NUMBER OF SEQ ID NOS: 70
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 52
;   LENGTH: 4952
    TYPE: DNA
    ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: pXL-Bac-EGFP
;   OTHER INFORMATION: sequence
US-10-001-189-52

  Query Match            28.2%;  Score 28.2;  DB 9;  Length 4952;
  Best Local Similarity  61.6%;  Pred. No. 3.8;
  Matches   45;  Conservative    0;  Mismatches   28;  Indels    0;  Gaps    0;

Qy      18 CGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGAACAGAAC 77
           | || || || | | |||  || | || |||||| || |    ||| ||| |||| |
Db     778 CTGATGCGCTTCGGGCTGACCATCCGGAACTGTGTCCGGAAAAGCCGCGACGAACTGGTA 837

Qy      78 TCTGTGGTGTGCT 90
           ||   ||||  ||
Db     838 TCCCAGGTGGCCT 850


RESULT 14
US-10-001-189-48
; Sequence 48, Application US/10001189
; Patent No. US20020173634A1
; GENERAL INFORMATION:
;  APPLICANT: FRASER JR., MALCOLM J.
;  APPLICANT:  LI, XU
;  APPLICANT:  BEAM, TERESA
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
   TITLE OF INVENTION:  MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
;  TITLE OF INVENTION:  VECTOR PIGGYBAC
;  FILE REFERENCE: 835910-92098
;  CURRENT APPLICATION NUMBER: US/10/001,189
;  CURRENT FILING DATE:  2001-10-30
;  PRIOR APPLICATION NUMBER: 60/244,984
;  PRIOR FILING DATE: 2000-11-01
;  PRIOR APPLICATION NUMBER: 60/244,677
;  PRIOR FILING DATE: 2000-10-31
;  NUMBER OF SEQ ID NOS: 70
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 48
;   LENGTH: 8999
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: p(PZ)-Bac-EYFP
;   OTHER INFORMATION: sequence
US-10-001-189-48

  Query Match            28.2%;  Score 28.2;  DB 9;  Length 8999;
  Best Local Similarity  61.6%;  Pred. No. 4.4;
  Matches   45;  Conservative    0;  Mismatches   28;  Indels    0;  Gaps    0;

Qy      18 CGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGAACAGAAC 77
           | || || || | | |||  || | || |||||| || |    ||| ||| |||| |
Db    8593 CTGATGCGCTTCGGGCTGACCATCCGGAACTGTGTCCGGAAAAGCCGCGACGAACTGGTA 8652

Qy      78 TCTGTGGTGTGCT 90
           ||   ||||  ||
Db    8653 TCCCAGGTGGCCT 8665


RESULT 15
US-10-001-189-49
; Sequence 49, Application US/10001189
; Patent No. US20020173634A1
; GENERAL INFORMATION:
;  APPLICANT: FRASER JR., MALCOLM J.
;  APPLICANT:  LI, XU
;  APPLICANT:  BEAM, TERESA
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
;  TITLE OF INVENTION:  MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
;  TITLE OF INVENTION:  VECTOR PIGGYBAC
;  FILE REFERENCE: 835910-92098
;  CURRENT APPLICATION NUMBER: US/10/001,189
;  CURRENT FILING DATE:  2001-10-30
;  PRIOR APPLICATION NUMBER: 60/244,984
;  PRIOR FILING DATE: 2000-11-01
;  PRIOR APPLICATION NUMBER: 60/244,677
;  PRIOR FILING DATE: 2000-10-31
;  NUMBER OF SEQ ID NOS: 70
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 49
;   LENGTH: 9012
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: p(PZ)-Bac-ECFP
;   OTHER INFORMATION: sequence
US-10-001-189-49

  Query Match            28.2%;  Score 28.2;  DB 9;  Length 9012;
  Best Local Similarity  61.6%;  Pred. No. 4.4;
  Matches   45;  Conservative    0;  Mismatches   28;  Indels    0;  Gaps    0;

Qy      18 CGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGAACAGAAC 77
           | || || || | | |||  || | || |||||| || |    ||| ||| |||| |
Db    8606 CTGATGCGCTTCGGGCTGACCATCCGGAACTGTGTCCGGAAAAGCCGCGACGAACTGGTA 8665

Qy      78 TCTGTGGTGTGCT 90
           ||   ||||  ||
Db    8666 TCCCAGGTGGCCT 8678


Search completed: January 15, 2003, 22:32:21
Job time : 36.5 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2003, 19:39:40 ; Search time 1268 Seconds
(without alignments)
1277.247 Million cell updates/sec

Title: L00394_COPY_1_100
Perfect score: 100
Sequence: 1 CTCCTTTGGCAGTCACACGG.....GTGGTGTGCTCCTGCGCCCG 100

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	88.4	88.4	407	12	BE839155	BE839155 RC6-FN013
	2	88.4	88.4	716	13	BI757668	BI757668 603028107
	3	88.4	88.4	844	9	AL553470	AL553470 AL553470
	4	88.4	88.4	882	14	BQ882177	BQ882177 AGENCOURT
	5	88.4	88.4	898	9	AL521984	AL521984 AL521984
	6	88.4	88.4	984	9	AL570383	AL570383 AL570383
	7	88.4	88.4	1049	13	BM546790	BM546790 AGENCOURT
	8	88.4	88.4	1211	14	BM926296	BM926296 AGENCOURT
	9	87.4	87.4	455	14	W21335	W21335 zb55f03.r1
c	10	86.8	86.8	472	12	BE838990	BE838990 RC6-FN013
	11	65.4	65.4	552	9	AI255604	AI255604 ui56a05.y
	12	65.4	65.4	596	13	BI328381	BI328381 602985811
	13	65.4	65.4	758	13	BI257017	BI257017 602978857
	14	65.4	65.4	771	12	BF384882	BF384882 602046381
	15	65.4	65.4	784	12	BF532104	BF532104 602073227
	16	65.4	65.4	875	12	BF385197	BF385197 602047517
	17	62.8	62.8	790	12	BF235332	BF235332 602026725
	18	53.6	53.6	698	13	BI146007	BI146007 602911359
	19	39	39.0	598	12	BG641612	BG641612 pg11c.pk0
	20	39	39.0	640	12	BG641885	BG641885 pg11c.pk0
	21	39	39.0	640	13	BI468182	BI468182 EST00589
	22	37.6	37.6	1061	17	CNS039FM	AL233851 Tetraodon
	23	36.4	36.4	638	10	AW133828	AW133828 fi12c10.y
	24	36.2	36.2	641	13	BM427276	BM427276 pgf2n.pk0
	25	35.6	35.6	900	13	BM438559	BM438559 IpLvr0030
	26	33	33.0	325	9	AA176034	AA176034 mt04e07.r
	27	33	33.0	433	10	BB796863	BB796863 BB796863
	28	33	33.0	687	9	AV285869	AV285869 AV285869
c	29	32.2	32.2	834	17	CNS02RM1	AL210754 Tetraodon
c	30	32	32.0	285	9	AI449301	AI449301 mt04e07.x
c	31	31.6	31.6	304	9	AI852041	AI852041 UI-M-BH0-
	32	31.4	31.4	349	10	BB798742	BB798742 BB798742
	33	31.4	31.4	359	10	BB804261	BB804261 BB804261
	34	31.4	31.4	385	10	BB806529	BB806529 BB806529
	35	31.4	31.4	956	12	BG122154	BG122154 602349596
c	36	30.8	30.8	206	12	BF912734	BF912734 PM1-UT010
c	37	30.8	30.8	295	9	AA293295	AA293295 zt28c10.s
	38	30.8	30.8	543	12	BG710842	BG710842 pg11n.pk0
c	39	30.8	30.8	602	14	BM951218	BM951218 UI-M-EH0-
c	40	30.8	30.8	679	14	BQ330929	BQ330929 MR0-ET010
	41	30.8	30.8	720	13	BI552439	BI552439 603196138
	42	30.8	30.8	724	13	BI547738	BI547738 603191892
	43	30.8	30.8	749	13	BM050541	BM050541 603632827
c	44	30.8	30.8	776	13	BI833415	BI833415 603088064
	45	30.8	30.8	848	13	BI752594	BI752594 603028432

ALIGNMENTS

RESULT 1
BE839155/c
LOCUS BE839155 407 bp mRNA linear EST 22-SEP-2000
DEFINITION RC6-FN0138-110800-012-E07 FN0138 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE839155
VERSION BE839155.1 GI:10271442
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 407)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001

```
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=RC6-FN0138-110
            800-012-E07&t3=2000-08-11&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 18
            High quality sequence stop: 352.
FEATURES             Location/Qualifiers
     source          1. .407
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="FN0138"
                     /dev_stage="Adult"
                     /note="Organ: prostate_normal; Vector: puc18; Site_1: SmaI
                     ; Site_2: SmaI; A mini-library was made by cloning
                     products derived from  ORESTES PCR (U.S. Letters Patent
                     application No. 196,716 -  Ludwig Institute for Cancer
                     Research) profiles into  the pUC 18 vector. Reverse
                     transcription of tissue mRNA and cDNA amplification were
                     performed under low stringency conditions."
 ASE COUNT       64 a    123 c    107 g    113 t
ORIGIN

  Query Match            88.4%;  Score 88.4;  DB 12;  Length 407;
  Best Local Similarity  98.9%;  Pred. No. 4.9e-16;
  Matches   89;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      11 AGTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 70
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     210 ATTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 151

Qy      71 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 100
           ||||||||||||||||||||||||||||||
Db     150 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 121


RESULT 2
BI757668
LOCUS       BI757668                716 bp    mRNA    linear   EST 25-SEP-2001
DEFINITION  603028107F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5198760 5',
            mRNA sequence.
ACCESSION   BI757668
VERSION     BI757668.1  GI:15749246
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  FERENCE   1  (bases 1 to 716)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11497  row: m  column: 01
            High quality sequence stop: 716.
FEATURES             Location/Qualifiers
     source          1. .716
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5198760"
                     /clone_lib="NIH_MGC_114"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: EcoRV (destroyed); RNA source anonymous pool of 6
                     male brains, age range 23-27 yo. Library is oligo-dT
                     primed and directionally cloned (EcoRV site is  destroyed
                     upon cloning).  Average insert size 1.5 kb, insert size
                     range 1-3 kb.  Library is normalized and enriched for
                     full-length  clones and was constructed by C. Gruber
                     (Invitrogen).  Research  Genetics tracking code 019. Note:
                     this is a NIH_MGC Library."
BASE COUNT      189 a    204 c    210 g    113 t
ORIGIN

  Query Match            88.4%;  Score 88.4;  DB 13;  Length 716;
  Best Local Similarity  98.9%;  Pred. No. 6.5e-16;
  Matches   89;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      11 AGTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 70
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     369 ATTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 428

Qy      71 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 100
           ||||||||||||||||||||||||||||||
Db     429 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 458


RESULT 3
AL553470
LOCUS       AL553470                844 bp    mRNA    linear   EST 16-FEB-2001
DEFINITION  AL553470 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0DI076YA10 5
            prime, mRNA sequence.
ACCESSION   AL553470
VERSION     AL553470.1  GI:12893334
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 844)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             Location/Qualifiers
     source          1. .844
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CS0DI076YA10"
                     /clone_lib="LTI_NFL006_PL2"
                     /tissue_type="placenta"
                     /note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer.  Five prime end
                     enriched, double-stranded cDNA was digested with Not I and
                     cloned into the Not I and Eco RV sites of the pCMVSPORT 6
                     vector. Library was normalized. Library was constructed by
                     Life Technologies. Contact : Feng Liang Life Technologies,
                     a division of Invitrogen 9800 Medical Center Drive
                     Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
                     Email : fliang@lifetech.com URL :
                     http://fulllength.invitrogen.com"
BASE COUNT      221 a    231 c    244 g    137 t     11 others
ORIGIN

  Query Match            88.4%;  Score 88.4;  DB 9;  Length 844;
  Best Local Similarity  98.9%;  Pred. No. 7e-16;
  Matches   89;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      11 AGTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 70
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     398 ATTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 457

Qy      71 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 100
           ||||||||||||||||||||||||||||||
```

```
Db    458 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 487


RESULT 4
BQ882177
LOCUS       BQ882177                882 bp    mRNA    linear   EST 16-AUG-2002
DEFINITION  AGENCOURT_7978442 Lupski_dorsal_root_ganglion Homo sapiens cDNA
            clone IMAGE:6185224 5', mRNA sequence.
ACCESSION   BQ882177
VERSION     BQ882177.1  GI:22274185
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 882)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. James R. Lupski
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13575  row: k  column: 17
            High quality sequence stop: 742.
FEATURES             Location/Qualifiers
     source          1. .882
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6185224"
                     /clone_lib="Lupski_dorsal_root_ganglion"
                     /sex="male"
                     /tissue_type="dorsal root ganglia"
                     /dev_stage="adult, 36 yr"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
                     NotI; Site_2: SalI; cDNA made by oligo-dT priming.
                     Directionally cloned using the following adaptors:
                     5'-TCGACCCACGCGTCCG-3' and
                     5'-GACTAGTTCTAGATCGCGAGCGGCCGCCCT(15)-3'. Size selected >
                     1 kb for average insert length 1.7 kb. This is a primary
                     library, non-amplified. Library constructed by Life
                     Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
                     College of Medicine) and is available through Life
                     Technologies."
BASE COUNT      225 a    250 c    262 g    144 t      1 others
ORIGIN

  Query Match             88.4%;  Score 88.4;  DB 14;  Length 882;
  Best Local Similarity   98.9%;  Pred. No. 7.2e-16;
  Matches   89;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      11 AGTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 70
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     433 ATTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 492

Qy      71 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 100
           |||||||||||||||||||||||||||||
Db     493 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 522


RESULT 5
AL521984
LOCUS       AL521984                898 bp    mRNA    linear   EST 13-FEB-2001
DEFINITION  AL521984 LTI_NFL004_NBC2 Homo sapiens cDNA clone CS0DB007YC01 5
            prime, mRNA sequence.
ACCESSION   AL521984
VERSION     AL521984.1  GI:12785477
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 898)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             Location/Qualifiers
     source          1. .898
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CS0DB007YC01"
                     /clone_lib="LTI_NFL004_NBC2"
                     /sex="male"
                     /tissue_type="neuroblastoma cells"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-stranded cDNA was digested with Not I and
                     cloned into the Not I and Eco RV sites of the pCMVSPORT 6
                     vector.  Library was normalized. Library was constructed
                     by Life Technologies.  Contact : Feng Liang Life
                     Technologies, a division of Invitrogen 9800 Medical Center
                     Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
                     8371 Email : fliang@lifetech.com URL :
                     http://fulllength.invitrogen.com"
BASE COUNT      234 a    253 c    264 g    147 t
ORIGIN

  Query Match             88.4%;  Score 88.4;  DB 9;  Length 898;
  Best Local Similarity   98.9%;  Pred. No. 7.2e-16;
  Matches   89;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      11 AGTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 70
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     389 ATTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 448

Qy      71 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 100
           |||||||||||||||||||||||||||||
Db     449 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 478


RESULT 6
AL570383
LOCUS       AL570383                984 bp    mRNA    linear   EST 16-FEB-2001
DEFINITION  AL570383 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0DI020YH13 5
            prime, mRNA sequence.
ACCESSION   AL570383
VERSION     AL570383.1  GI:12926636
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 984)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             Location/Qualifiers
     source          1. .984
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CS0DI020YH13"
```

```
                     /clone_lib="LTI_NFL006_PL2"
                     /tissue_type="placenta"
                     /note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer.  Five prime end
                     enriched, double-stranded cDNA was digested with Not I and
                     cloned into the Not I and Eco RV sites of the pCMVSPORT 6
                     vector. Library was normalized. Library was constructed by
                     Life Technologies. Contact : Feng Liang Life Technologies,
                     a division of Invitrogen 9800 Medical Center Drive
                     Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
                     Email : fliang@lifetech.com URL :
                     http://fulllength.invitrogen.com"
BASE COUNT      252 a    269 c    300 g    160 t      3 others
ORIGIN

  Query Match             88.4%;  Score 88.4;  DB 9;  Length 984;
  Best Local Similarity   98.9%;  Pred. No. 7.5e-16;
  Matches   89;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      11 AGTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 70
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     432 ATTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 491

Qy      71 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 100
           ||||||||||||||||||||||||||||||
Db     492 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 521


RESULT 7
BM546790
LOCUS       BM546790                1049 bp    mRNA    linear   EST 20-FEB-2002
DEFINITION  AGENCOURT_6491143 NIH_MGC_125 Homo sapiens cDNA clone IMAGE:5723510
            5', mRNA sequence.
ACCESSION   BM546790
VERSION     BM546790.1  GI:18780039
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1049)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Invitrogen
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12711  row: e  column: 15
            High quality sequence stop: 759.
FEATURES             Location/Qualifiers
     source          1. .1049
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5723510"
                     /clone_lib="NIH_MGC_125"
                     /lab_host="DH10B"
                     /note="Organ: ovary (pool of 3); Vector: pCMV-SPORT6;
                     Site_1: EcoRV (destroyed); Site_2: NotI; RNA source pool
                     of three ovaries, from females ranging in age from 38 to
                     49 yo. Library is oligo-dT primed and directionally cloned
                     (EcoRV site is destroyed upon cloning). Average insert
                     size 2.1 kb, insert size range 1-3.5 kb. Library is
                     normalized and enriched for full-length clones and was
                     constructed by C. Gruber (Invitrogen). Research Genetics
                     tracking code 036."
BASE COUNT      274 a    289 c    316 g    169 t      1 others
ORIGIN

  Query Match             88.4%;  Score 88.4;  DB 13;  Length 1049;
  Best Local Similarity   98.9%;  Pred. No. 7.8e-16;
  Matches   89;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      11 AGTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 70
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     426 ATTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 485

Qy      71 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 100
           ||||||||||||||||||||||||||||||
Db     486 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 515


RESULT 8
BM926296
LOCUS       BM926296                1211 bp    mRNA    linear   EST 12-MAR-2002
DEFINITION  AGENCOURT_6600735 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5764698
            5', mRNA sequence.
ACCESSION   BM926296
VERSION     BM926296.1  GI:19376675
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1211)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12818  row: i  column: 19
            High quality sequence stop: 529.
FEATURES             Location/Qualifiers
     source          1. .1211
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5764698"
                     /clone_lib="NIH_MGC_114"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: EcoRV (destroyed); RNA source anonymous pool of 6
                     male brains, age range 23-27 yo. Library is oligo-dT
                     primed and directionally cloned (EcoRV site is  destroyed
                     upon cloning).  Average insert size 1.5 kb, insert size
                     range 1-3 kb.  Library is normalized and enriched for
                     full-length  clones and was constructed by C. Gruber
                     (Invitrogen).  Research  Genetics tracking code 019. Note:
                     this is a NIH_MGC Library."
BASE COUNT      296 a    351 c    351 g    211 t      2 others
ORIGIN

  Query Match             88.4%;  Score 88.4;  DB 14;  Length 1211;
  Best Local Similarity   98.9%;  Pred. No. 8.3e-16;
  Matches   89;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      11 AGTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 70
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     392 ATTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 451

Qy      71 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 100
           ||||||||||||||||||||||||||||||
Db     452 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 481
```

```
RESULT 9
W21335
LOCUS       W21335                  455 bp    mRNA    linear   EST 20-AUG-1996
DEFINITION  zb55f03.r1 Soares_fetal_lung_NbHL19W Homo sapiens cDNA clone
            IMAGE:307517 5' similar to gb:M57285 COAGULATION FACTOR X PRECURSOR
            (HUMAN);, mRNA sequence.
ACCESSION   W21335
VERSION     W21335.1  GI:1298387
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 455)
  AUTHORS   Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
            Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins
            ,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore
            ,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T.,
            Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J., Trevaskis,E.,
            Underwood,K., Wohldmann,P., Waterston,R., Wilson,R. and Marra,M.
  TITLE     Generation and analysis of 280,000 human expressed sequence tags
  JOURNAL   Genome Res. 6 (9), 807-828 (1996)
  MEDLINE   97044478
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1028   Std Error: 0.00
            Seq primer: mob.REGA+ET
            High quality sequence stop: 424.
FEATURES             Location/Qualifiers
     source          1. .455
                     /organism="Homo sapiens"
                     /db_xref="GDB:1250929"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:307517"
                     /clone_lib="Soares_fetal_lung_NbHL19W"
                     /dev_stage="19 weeks"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: lung; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer
                     [5'-TGTTACCAATCTGAAGTGGGAGCGGCCGCAATTTTTTTTTTTTTTTT-3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library went through one round of
                     normalization to a Cot = 5. Library constructed by Bento
                     Soares and M.Fatima Bonaldo. This library was constructed
                     from the same fetus as the fetal heart library, Soares
                     fetal heart NbHH19W."
BASE COUNT      116 a    135 c    131 g     69 t      4 others
ORIGIN

  Query Match             87.4%;  Score 87.4;  DB 14;  Length 455;
  Best Local Similarity   98.9%;  Pred. No. 1e-15;
  Matches   88;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      11 AGTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 70
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      96 ATTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 155

Qy      71 ACAGAACTCTGTGGTGTGCTCCTGCGCCC 99
           |||||||||||||||||||||||||||||
Db     156 ACAGAACTCTGTGGTGTGCTCCTGCGCCC 184


RESULT 10
BE838990/c
LOCUS       BE838990                472 bp    mRNA    linear   EST 22-SEP-2000
DEFINITION  RC6-FN0138-260700-011-G08 FN0138 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BE838990
VERSION     BE838990.1  GI:10271368
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 472)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=RC6-FN0138-260
            700-011-G08&t3=2000-07-26&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 9
            High quality sequence stop: 307.
FEATURES             Location/Qualifiers
     source          1. .472
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="FN0138"
                     /dev_stage="Adult"
                     /note="Organ: prostate_normal; Vector: puc18; Site_1: SmaI
                     ; Site_2: SmaI; A mini-library was made by cloning
                     products derived from  ORESTES PCR (U.S. Letters Patent
                     application No. 196,716 -  Ludwig Institute for Cancer
                     Research) profiles into  the pUC 18 vector. Reverse
                     transcription of tissue mRNA and cDNA amplification were
                     performed under low stringency conditions."
BASE COUNT      105 a    144 c    106 g    117 t
ORIGIN

  Query Match             86.8%;  Score 86.8;  DB 12;  Length 472;
  Best Local Similarity   97.8%;  Pred. No. 1.6e-15;
  Matches   88;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      11 AGTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 70
           | ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db     192 ATTCACACGGAAGCTCTGCGGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 133

Qy      71 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 100
           ||||||||||||||||||||||||||||||
Db     132 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 103


RESULT 11
AI255604
LOCUS       AI255604                552 bp    mRNA    linear   EST 12-NOV-1998
DEFINITION  ui56a05.y1 Sugano mouse liver mlia Mus musculus cDNA clone
            IMAGE:1886384 5' similar to gb:M57285 COAGULATION FACTOR X
            PRECURSOR (HUMAN);, mRNA sequence.
ACCESSION   AI255604
VERSION     AI255604.1  GI:3863129
KEYWORDS    EST.
```

```
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 552)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further informa'ion.
            MGI:970708
            Seq primer: custom primer used
            High quality sequence stop: 511.
FEATURES             Location/Qualifiers
     source          1. .552
                     /organism="Mus musculus"
                     /strain="C57BL"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1886384"
                     /clone_lib="Sugano mouse liver mlia"
                     /sex="female"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Organ: liver; Vector: pME18S-FL3; Site_1: DraIII
                     (CACTGTGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA
                     was primed with an oligo(dT) primer
                     [ATGTGGCCTTTTTTTTTTTTTTTTT]; double-stranded cDNA was
                     ligated to a  DraIII adaptor [TGTTGGCCTACTGG], digested
                     and cloned into distinct DraIII sites of the pME18S-FL3
                     vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should
                     be used to isolate the cDNA insert.  Size selection was
                     performed to exclude fragments <1.5kb.  Library
                     constructed by Dr. Sumio Sugano (University of Tokyo
                     Institute of Medical Science).  Custom primers  for
                     sequencing: 5' end primer CTTCTGCTCTAAAAGCTGCG and 3' end
                     primer CGACCTGCAGCTCGAGCACA."
BASE COUNT      147 a    129 c    156 g    120 t
 RIGIN

  Query Match             65.4%;  Score 65.4;  DB 9;  Length 552;
  Best Local Similarity   86.7%;  Pred. No. 3.7e-09;
  Matches   72;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy      18 CGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGAACAGAAC 77
           ||||| |||||| |||| |||||||| ||||||||||||||||||    || || ||||||
Db     390 CGGAAACTCTGCCGCCTAGACAACGGAGACTGTGACCAGTTCTGCAGAGAAGAGCAGAAC 449

Qy      78 TCTGTGGTGTGCTCCTGCGCCCG 100
           || |||||||||||||||||| |
Db     450 TCAGTGGTGTGCTCCTGCGCCAG 472


RESULT 12
BI328381
LOCUS       BI328381                596 bp    mRNA    linear   EST 30-JUL-2001
DEFINITION  602985811F1 NCI_CGAP_Li9 Mus musculus cDNA clone IMAGE:5138306 5',
            mRNA sequence.
ACCESSION   BI328381
VERSION     BI328381.1  GI:15013038
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 596)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey E. Green, M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11340  row: f  column: 03
            High quality sequence stop: 594.
FEATURES             Location/Qualifiers
     source          1. .596
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5138306"
                     /clone_lib="NCI_CGAP_Li9"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /note="Organ: liver; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                     Average insert size 1.9 kb. Constructed by Life
                     Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT      159 a    140 c    169 g    128 t
ORIGIN

  Query Match             65.4%;  Score 65.4;  DB 13;  Length 596;
  Best Local Similarity   86.7%;  Pred. No. 3.8e-09;
  Matches   72;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy      18 CGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGAACAGAAC 77
           ||||| |||||| |||| |||||||| ||||||||||||||||||    || || ||||||
Db     384 CGGAAACTCTGCCGCCTAGACAACGGAGACTGTGACCAGTTCTGCAGAGAAGAGCAGAAC 443

Qy      78 TCTGTGGTGTGCTCCTGCGCCCG 100
           || |||||||||||||||||| |
Db     444 TCAGTGGTGTGCTCCTGCGCCAG 466


RESULT 13
BI257017
LOCUS       BI257017                758 bp    mRNA    linear   EST 20-JUL-2001
DEFINITION  602978857F1 NCI_CGAP_Li9 Mus musculus cDNA clone IMAGE:5123607 5',
            mRNA sequence.
ACCESSION   BI257017
VERSION     BI257017.1  GI:14811984
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 758)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: Incyte Genomics, Inc.
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11302  row: a  column: 16
            High quality sequence start: 3
            High quality sequence stop: 756.
```

```
FEATURES             Location/Qualifiers
     source          1. .758
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5123607"
                     /clone_lib="NCI_CGAP_Li9"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /note="Organ: liver; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                     Average insert size 1.9 kb. Constructed by Life
                     Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT      200 a    171 c    223 g    164 t
ORIGIN

  Query Match             65.4%;  Score 65.4;  DB 13;  Length 758;
  Best Local Similarity   86.7%;  Pred. No. 4.3e-09;
  Matches   72;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy      18 CGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGAACAGAAC 77
           ||||| |||||| |||| |||||||| |||||||||||||||||   || || ||||||
Db     378 CGGAAACTCTGCCGCCTAGACAACGGAGACTGTGACCAGTTCTGCAGAGAAGAGCAGAAC 437

Qy      78 TCTGTGGTGTGCTCCTGCGCCCG 100
           || |||||||||||||||||| |
Db     438 TCAGTGGTGTGCTCCTGCGCCAG 460


RESULT 14
BF384882
LOCUS       BF384882                771 bp    mRNA    linear   EST 27-NOV-2000
DEFINITION  602046381F1 NCI_CGAP_Li9 Mus musculus cDNA clone IMAGE:4195508 5',
            mRNA sequence.
ACCESSION   BF384882
VERSION     BF384882.1  GI:11366187
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 771)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey E. Green, M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9531  row: b  column: 21
            High quality sequence stop: 690.
FEATURES             Location/Qualifiers
     source          1. .771
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:4195508"
                     /clone_lib="NCI_CGAP_Li9"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /note="Organ: liver; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                     Average insert size 1.9 kb. Constructed by Life
                     Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT      204 a    184 c    231 g    152 t
ORIGIN

  Query Match             65.4%;  Score 65.4;  DB 12;  Length 771;
  Best Local Similarity   86.7%;  Pred. No. 4.4e-09;
  Matches   72;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy      18 CGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGAACAGAAC 77
           ||||| |||||| |||| |||||||| |||||||||||||||||   || || ||||||
Db     381 CGGAAACTCTGCCGCCTAGACAACGGAGACTGTGACCAGTTCTGCAGAGAAGAGCAGAAC 440

Qy      78 TCTGTGGTGTGCTCCTGCGCCCG 100
           || |||||||||||||||||| |
Db     441 TCAGTGGTGTGCTCCTGCGCCAG 463


RESULT 15
BF532104
LOCUS       BF532104                784 bp    mRNA    linear   EST 11-DEC-2000
DEFINITION  602073227F1 NCI_CGAP_Li9 Mus musculus cDNA clone IMAGE:4210183 5',
            mRNA sequence.
ACCESSION   BF532104
VERSION     BF532104.1  GI:11619570
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 784)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey E. Green, M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9777  row: f  column: 08
            High quality sequence stop: 722.
FEATURES             Location/Qualifiers
     source          1. .784
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:4210183"
                     /clone_lib="NCI_CGAP_Li9"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /note="Organ: liver; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                     Average insert size 1.9 kb. Constructed by Life
                     Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT      202 a    186 c    231 g    165 t
ORIGIN

  Query Match             65.4%;  Score 65.4;  DB 12;  Length 784;
  Best Local Similarity   86.7%;  Pred. No. 4.4e-09;
  Matches   72;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy      18 CGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGAACAGAAC 77
           ||||| |||||| |||| |||||||| |||||||||||||||||   || || ||||||
Db     383 CGGAAACTCTGCCGCCTAGACAACGGAGACTGTGACCAGTTCTGCAGAGAAGAGCAGAAC 442

Qy      78 TCTGTGGTGTGCTCCTGCGCCCG 100
           || |||||||||||||||||| |
Db     443 TCAGTGGTGTGCTCCTGCGCCAG 465


Search completed: January 15, 2003, 21:18:47
Job time : 1273 secs
```

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: January 15, 2003, 17:50:20 ; Search time 151 Seconds
(without alignments)
1491.390 Million cell updates/sec

Title: L00394_COPY_1_100
Perfect score: 100
Sequence: 1 CTCCTTTGGCAGTCACACGG.....GTGGTGTGCTCCTGCGCCCG 100

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	100	100.0	367	21	AAC70860	Single nucleotide
	2	99.6	99.6	367	21	AAC70878	Single nucleotide
	3	99.6	99.6	367	21	AAC70884	Single nucleotide
	4	88.4	88.4	788	12	AAQ12776	Factor X-LACI hybr
	5	88.4	88.4	1126	20	AAX15427	DNA encoding coagu
	6	88.4	88.4	1126	21	AAA89786	DNA encoding coagu
	7	88.4	88.4	1126	21	AAA12970	DNA encoding Facto
	8	88.4	88.4	1126	21	AAZ56120	Vitamin-K-dependen
	9	88.4	88.4	1404	19	AAV10462	Human Factor X pro
	10	88.4	88.4	1467	19	AAV56776	Human Factor X gen
	11	88.4	88.4	1467	19	AAV56821	Human Factor X gen
	12	88.4	88.4	1467	22	AAF59409	Human factor X nuc
	13	88.4	88.4	1507	21	AAA54031	Human factor X cod
	14	88.4	88.4	1887	22	AAH57469	Human liver cell s
	15	85.4	85.4	300	22	AAH57261	Human liver specif
	16	82.8	82.8	1560	22	AAF24735	Nucleotide sequenc
	17	82.8	82.8	1860	22	AAF24738	Nucleotide sequenc
	18	52.4	52.4	1554	15	AAQ71243	Serine protease fo
	19	31.8	31.8	13923	23	ABL05109	Drosophila melanog
c	20	31.8	31.8	17902	23	ABL05108	Drosophila melanog
c	21	30.8	30.8	295	24	ABL80210	Human ovarian canc
	22	30.8	30.8	2461	17	AAT41544	Human gas6-encodin
	23	29.6	29.6	4790	22	AAS45254	cDNA encoding nove
	24	29.6	29.6	7037	22	AAS45066	cDNA encoding nove
	25	29.4	29.4	4181	22	AAD06778	Human haematopoiet
	26	29.4	29.4	4801	22	AAD06781	Human haematopoiet
c	27	29.2	29.2	1743	23	AAS85732	DNA encoding novel
	28	29.2	29.2	2196	23	AAS82534	DNA encoding novel
c	29	29.2	29.2	2196	23	AAS85733	DNA encoding novel
	30	29.2	29.2	2196	23	AAS89351	DNA encoding novel
	31	29.2	29.2	2909	23	AAS73293	DNA encoding novel
c	32	29.2	29.2	3747	23	AAS73044	DNA encoding novel
	33	29.2	29.2	3810	23	AAS82533	DNA encoding novel
	34	29.2	29.2	3810	23	AAS89350	DNA encoding novel
c	35	29.2	29.2	4901	23	AAS92853	DNA encoding novel
c	36	29.2	29.2	6813	23	AAS85740	DNA encoding novel
	37	29	29.0	1241	23	AAS77139	DNA encoding novel
	38	29	29.0	2013	22	AAL03051	Human reproductive
	39	29	29.0	2013	23	ABL97387	Human testicular a
	40	29	29.0	2565	24	ABL92075	Human Tumour Endot
	41	29	29.0	2565	24	ABL92086	Human Tumour Endot
c	42	28.6	28.6	405	22	AAL01005	Human reproductive
c	43	28.6	28.6	405	23	ABL96473	Human testicular a
	44	28.6	28.6	719	23	ABL05065	Drosophila melanog
c	45	28.6	28.6	1875	23	AAS85741	DNA encoding novel

ALIGNMENTS

RESULT 1
AAC70860
ID AAC70860 standard; DNA; 367 BP.
XX
AC AAC70860;
XX
DT 09-FEB-2001 (first entry)
XX
DE Single nucleotide polymorphism containing sequence #230.
XX
KW Single nucleotide polymorphism; SNP; human; genetic disease;
KW disease susceptibility; cardiovascular system; endocrine system;
KW neurological system; forensic testing; paternity testing; ds.
XX
OS Homo sapiens.
XX
PN WO200058519-A2.
XX
PD 05-OCT-2000.
XX
PF 30-MAR-2000; 2000WO-US08440.
XX
PR 31-MAR-1999; 99US-0127248.
XX
PA (WHED) WHITEHEAD INST BIOMEDICAL RES.
PA (AFFY-) AFFYMETRIX INC.
XX
PI Altshuler D, Cargill M, Daley GQ, Ireland JS, Lander ES;
PI Lipshutz RJ, Patil N, Sklar P;
XX
DR WPI; 2000-611722/58.
XX

```
PT   Nucleic acid selected from one of 106 genes comprising single
PT   nucleotide polymorphisms, allele-specific oligonucleotides to the genes
PT   are useful for phenotypic correlations, forensics, paternity testing,
PT   medicine and genetic analysis -
XX
PS   Claim 1; Fig 5; 214pp; English.
XX
CC   The present invention is concerned with a number of human single
CC   nucleotide polymorphisms (SNPs) which the inventors identified in human
CC   genes. These SNPs can be used in disease diagnosis and prediction of an
CC   individual's susceptibility to disease, in forensic and paternity testing
CC   and in genetic mapping. In particular, the SNPs of the invention can be
CC   used to diagnose susceptibility to diseases of the cardiovascular,
CC   endocrine and neurological systems, such as coronary artery disease,
CC   schizophrenia, cancer, autoimmune diseases, Alzheimer's and Parkinson's
CC   diseases.
CC   Note: The degenerate codon within the sequence represents the position
CC   of an SNP, for example the letter S represents a polymorphism where the
CC   nucleotide may be C or G.
XX
SQ   Sequence 367 BP; 66 A; 122 C; 103 G; 75 T; 1 other;

  Query Match             100.0%;  Score 100;  DB 21;  Length 367;
  Best Local Similarity   100.0%;  Pred. No. 3.1e-22;
  Matches  100;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CTCCTTTGGCAGTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCT 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     212 CTCCTTTGGCAGTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCT 271

Qy      61 GCCACGAGGAACAGAACTCTGTGGTGTGCTCCTGCGCCCG 100
           ||||||||||||||||||||||||||||||||||||||||
Db     272 GCCACGAGGAACAGAACTCTGTGGTGTGCTCCTGCGCCCG 311


RESULT 2
AAC70878
ID   AAC70878 standard; DNA; 367 BP.
XX
AC   AAC70878;
XX
DT   09-FEB-2001  (first entry)
XX
DE   Single nucleotide polymorphism containing sequence #236.
XX
KW   Single nucleotide polymorphism; SNP; human; genetic disease;
KW   disease susceptibility; cardiovascular system; endocrine system;
KW   neurological system; forensic testing; paternity testing; ds.
XX
OS   Homo sapiens.
XX
PN   WO200058519-A2.
XX
PD   05-OCT-2000.
XX
PF   30-MAR-2000; 2000WO-US08440.
XX
PR   31-MAR-1999;   99US-0127248.
XX
PA   (WHED ) WHITEHEAD INST BIOMEDICAL RES.
PA   (AFFY-) AFFYMETRIX INC.
XX
PI   Altshuler D,  Cargill M,  Daley GQ,  Ireland JS,  Lander ES;
PI   Lipshutz RJ,  Patil N,  Sklar P;
XX
DR   WPI; 2000-611722/58.
XX
PT   Nucleic acid selected from one of 106 genes comprising single
PT   nucleotide polymorphisms, allele-specific oligonucleotides to the genes
PT   are useful for phenotypic correlations, forensics, paternity testing,
PT   medicine and genetic analysis -
XX
PS   Claim 1; Fig 5; 214pp; English.
XX
CC   The present invention is concerned with a number of human single
CC   nucleotide polymorphisms (SNPs) which the inventors identified in human
CC   genes. These SNPs can be used in disease diagnosis and prediction of an
CC   individual's susceptibility to disease, in forensic and paternity testing
CC   and in genetic mapping. In particular, the SNPs of the invention can be
CC   used to diagnose susceptibility to diseases of the cardiovascular,
CC   endocrine and neurological systems, such as coronary artery disease,
CC   schizophrenia, cancer, autoimmune diseases, Alzheimer's and Parkinson's
CC   diseases.
CC   Note: The degenerate codon within the sequence represents the position
CC   of an SNP, for example the letter S represents a polymorphism where the
CC   nucleotide may be C or G.
XX
SQ   Sequence 367 BP; 67 A; 121 C; 103 G; 75 T; 1 other;

  Query Match             99.6%;  Score 99.6;  DB 21;  Length 367;
  Best Local Similarity   99.0%;  Pred. No. 4.1e-22;
  Matches   99;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CTCCTTTGGCAGTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCT 60
           ||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||
Db     212 CTCCTTTGGCAGTCACACGGAAGCTCTGCAGCCTGGACAAYGGGGACTGTGACCAGTTCT 271

Qy      61 GCCACGAGGAACAGAACTCTGTGGTGTGCTCCTGCGCCCG 100
           ||||||||||||||||||||||||||||||||||||||||
Db     272 GCCACGAGGAACAGAACTCTGTGGTGTGCTCCTGCGCCCG 311


RESULT 3
AAC70884
ID   AAC70884 standard; DNA; 367 BP.
XX
AC   AAC70884;
XX
DT   09-FEB-2001  (first entry)
XX
DE   Single nucleotide polymorphism containing sequence #238.
XX
KW   Single nucleotide polymorphism; SNP; human; genetic disease;
KW   disease susceptibility; cardiovascular system; endocrine system;
KW   neurological system; forensic testing; paternity testing; ds.
XX
OS   Homo sapiens.
XX
PN   WO200058519-A2.
XX
PD   05-OCT-2000.
XX
PF   30-MAR-2000; 2000WO-US08440.
XX
PR   31-MAR-1999;   99US-0127248.
XX
PA   (WHED ) WHITEHEAD INST BIOMEDICAL RES.
PA   (AFFY-) AFFYMETRIX INC.
XX
PI   Altshuler D,  Cargill M,  Daley GQ,  Ireland JS,  Lander ES;
PI   Lipshutz RJ,  Patil N,  Sklar P;
XX
DR   WPI; 2000-611722/58.
XX
PT   Nucleic acid selected from one of 106 genes comprising single
PT   nucleotide polymorphisms, allele-specific oligonucleotides to the genes
PT   are useful for phenotypic correlations, forensics, paternity testing,
PT   medicine and genetic analysis -
XX
PS   Claim 1; Fig 5; 214pp; English.
XX
CC   The present invention is concerned with a number of human single
CC   nucleotide polymorphisms (SNPs) which the inventors identified in human
CC   genes. These SNPs can be used in disease diagnosis and prediction of an
CC   individual's susceptibility to disease, in forensic and paternity testing
CC   and in genetic mapping. In particular, the SNPs of the invention can be
```

```
CC   used to diagnose susceptibility to diseases of the cardiovascular,
CC   endocrine and neurological systems, such as coronary artery disease,
CC   schizophrenia, cancer, autoimmune diseases, Alzheimer's and Parkinson's
CC   diseases.
CC   Note: The degenerate codon within the sequence represents the position
CC   of an SNP, for example the letter S represents a polymorphism where the
CC   nucleotide may be C or G.
XX
SQ   Sequence 367 BP; 66 A; 122 C; 103 G; 75 T; 1 other;

  Query Match             99.6%;  Score 99.6;  DB 21;  Length 367;
  Best Local Similarity   99.0%;  Pred. No. 4.1e-22;
  Matches   99;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CTCCTTTGGCAGTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCT 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||
Db     212 CTCCTTTGGCAGTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCWGTTCT 271

Qy      61 GCCACGAGGAACAGAACTCTGTGGTGTGCTCCTGCGCCCG 100
           ||||||||||||||||||||||||||||||||||||||||
Db     272 GCCACGAGGAACAGAACTCTGTGGTGTGCTCCTGCGCCCG 311


RESULT 4
AAQ12776
ID   AAQ12776 standard; DNA; 788 BP.
XX
AC   AAQ12776;
XX
DT   09-OCT-1991  (first entry)
XX
DE   Factor X-LACI hybrid protein.
XX
KW   Kunitz domain; blood; coagulation; inhibitor; Factor X;
KW   tissue factor; TF; Lipoprotein-Associated Coagulation Inhibitor; ss.
XX
FH   Key             Location/Qualifiers
FT   CDS             22..771
FT                   /*tag= a
FT                   /label= X-LACI
FT   misc_RNA        22..534
FT                   /*tag= b
FT                   label= Xlc
FT   misc_RNA        535..771
FT                   /*tag= c
FT                   /label= kunitz_domain
FT   sig_peptide     22..141
FT                   /*tag= d
FT                   /label= pre-pro_leader
FT   misc_RNA        184..213
FT                   /*tag= e
FT                   /label= GLA_domain
FT   misc_RNA        285..471
FT                   /*tag= f
FT                   /label= growth_factor_domains
XX
PN   EP439442-A.
XX
PD   31-JUL-1991.
XX
PF   21-JAN-1991;   91EP-0870008.
XX
PR   25-JAN-1990;   90US-0470289.
XX
PA   (UNIW ) UNIV OF WASHINGTON.
XX
PI   Girard TJ,  Broze GJ;
XX
DR   WPI; 1991-224839/31.
DR   P-PSDB; AAR13675.
XX
PT   New factor X-LACI hybrid protein - comprises light chain of
PT   factor X and LACI's first kunitz domain for use as anticoagulant
XX
PS   Disclosure; Page 12-14; 17pp; English.
XX
CC   The protein is used as a blood coagulation inhibitor in mammals.  It
CC   is believed to mimic the Xa/LACI complex in binding to and
CC   inhibiting VIIa/tissue factor.  LACI inhibits via a novel feedback
CC   mechanism requiring generation of Xa (a prod. of VIIa/TF activity);
CC   XlcLACIki inhibits VIIa/TF activity directly.
CC   The DNA allows prodn. of XlcLACIki by introduction of the gene into
CC   cells suitable for expression, e.g. E. coli or CHO cells.
XX
SQ   Sequence 788 BP; 230 A; 187 C; 209 G; 162 T; 0 other;

  Query Match             88.4%;  Score 88.4;  DB 12;  Length 788;
  Best Local Similarity   98.9%;  Pred. No. 1.7e-18;
  Matches   89;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      11 AGTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 70
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     390 ATTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 449

Qy      71 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 100
           ||||||||||||||||||||||||||||||
Db     450 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 479


RESULT 5
AAX15427
ID   AAX15427 standard; DNA; 1126 BP.
XX
AC   AAX15427;
XX
DT   05-MAY-1999  (first entry)
XX
DE   DNA encoding coagulation factor X/Xa.
XX
KW   Truncated tissue factor; tissue factor binding ligand; coagulation;
KW   disease-associated vasculature; tumour; benign prostatic hyperplasia;
KW   diabetic-retinopathy; vascular restenosis; arteriovenous malformation;
KW   AVM; meningioma; hemangioma; neovascular glaucoma; psoriasis; synovitis;
KW   dermatitis; endometriosis; angiofibroma; rheumatoid arthritis;
KW   atheroscelotic plaque; corneal graft neovascularisation;
KW   haemophilic joint; hypertrophic scar; Osler-Weber syndrome;
KW   pyogenic granuloma retrolental fibroplasia; scleroderma; trachoma;
KW   vascular adhesion; coagulation factor; factor X/Xa; ss.
XX
OS   Homo sapiens.
XX
PN   US5877289-A.
XX
PD   02-MAR-1999.
XX
PF   07-JUN-1995;   95US-0479733.
XX
PR   07-JUN-1995;   95US-0479733.
PR   05-MAR-1992;   92US-0846349.
PR   02-MAR-1994;   94US-0205330.
PR   11-JUL-1994;   94US-0273567.
XX
PA   (SCRI ) SCRIPPS RES INST.
PA   (TEXA ) UNIV TEXAS SYSTEM.
XX
PI   Edgington TS,  Thorpe PE;
XX
DR   WPI; 1999-189722/16.
XX
PT   Tissue factor binding ligands  - comprising first binding region
PT   which binds to vasculature, particularly of tumours, and tissue
PT   factor construct
XX
PS   Example 9; Columns 129-132; 83pp; English.
XX
CC   The present sequence encodes a coagulation factor. The specification
```

```
CC   describes tissue factor binding ligands which comprise a binding region
CC   which binds to vasculature, particularly of tumours, and a tissue factor
CC   construct. The binding ligands can be used for stimulating coagulation in
CC   disease-associated vasculature, particularly for the treatment of
CC   tumours. The products can also be used for treating e.g. benign prostatic
CC   hyperplasia, diabetic-retinopathy, vascular restenosis, arteriovenous
CC   malformations (AVM), meningioma, hemangioma, neovascular glaucoma,
CC   psoriasis, synovitis, dermatitis, endometriosis, angiofibroma, rheumatoid
CC   arthritis, atheroscelotic plaques, corneal graft neovascularisation,
CC   haemophilic joints, hypertrophic scars, Osler-Weber syndrome, pyogenic
CC   granuloma retrolental fibroplasia, scleroderma, trachoma, or vascular
CC   adhesions. The products can also be used in binding assays.
XX
SQ   Sequence 1126 BP; 269 A; 341 C; 342 G; 174 T; 0 other;

  Query Match             88.4%;  Score 88.4;  DB 20;  Length 1126;
  Best Local Similarity   98.9%;  Pred. No. 1.9e-18;
  Matches   89;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      11 AGTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 70
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      27 ATTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 86

Qy      71 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 100
           ||||||||||||||||||||||||||||||
Db      87 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 116


RESULT 6
AAA89786
ID   AAA89786 standard; DNA; 1126 BP.
XX
AC   AAA89786;
XX
DT   14-DEC-2000  (first entry)
XX
DE   DNA encoding coagulation factor X/Xa.
XX
KW   Tissue factor protein; truncated tissue factor; tTF; cytostatic;
KW   coagulant; diabetic retinopathy; arteriovenous malformation;
KW   meningioma; hemangioma; neovascular glaucoma; psoriasis; synovitis;
KW   endometriosis; hemophylic joint; hypertrophic scar; vascular adhesion;
KW   tumour; cancer; ligand; human; factor X; ds.
XX
OS   Homo sapiens.
XX
PN   US6093399-A.
XX
     25-JUL-2000.

     07-JUN-1995;   95US-0482369.
XX
PR   05-MAR-1992;   92US-0846349.
PR   02-MAR-1994;   94US-0205330.
PR   11-JUL-1994;   94US-0273567.
XX
PA   (SCRI ) SCRIPPS RES INST.
PA   (TEXA ) UNIV TEXAS SYSTEM.
XX
PI   Edgington TS,  Thorpe PE;
XX
DR   WPI; 2000-531471/48.
XX
PT   New immunological and growth factor-based bispecific binding ligands,
PT   useful for stimulating coagulation in vasculature-associated diseases,
PT   e.g. for treating both benign and malignant diseases (e.g. meningioma
PT   or hemangioma)  -
XX
PS   Example 9; Column 129-130; 83pp; English.
XX
CC   The present invention relates to a binding ligand with a first binding
CC   region that is operatively linked to either a coagulation factor or a
CC   second binding region that binds to a coagulation factor. The first
CC   binding region binds to a component on the surface of a tumour. The
CC   second binding region is all or part of an antibody. An example of a
CC   coagulation factor for use in the invention is human truncated tissue
CC   factor. Truncated tissue factor (tTF) is the extracellular domain of the
CC   mature tissue factor protein (see AAB15019). The binding ligand of the
CC   invention is useful for stimulating coagulation in vasculature
CC   associated diseases. Particularly, the binding ligand is useful for
CC   treating both benign and malignant diseases that have a vascular
CC   component. These diseases include benign growths (e.g. BPH), diabetic
CC   retinopathy, arteriovenous malformations, meningioma, hemangioma,
CC   neovascular glaucoma, psoriasis, synovitis, endometriosis, hemophylic
CC   joints, hypertrophic scars or vascular adhesions. The present binding
CC   ligands offer the advantage that even limited damage to the tumour
CC   vasculature could produce an avalanche of tumour cell death because
CC   each capillary provides oxygen and nutrients for thousands of tumour
CC   cells. The present sequence is DNA encoding coagulation factor
CC   X/Xa. This factor was used in the invention.
XX
SQ   Sequence 1126 BP; 269 A; 341 C; 342 G; 174 T; 0 other;

  Query Match             88.4%;  Score 88.4;  DB 21;  Length 1126;
  Best Local Similarity   98.9%;  Pred. No. 1.9e-18;
  Matches   89;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      11 AGTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 70
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      27 ATTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 86

Qy      71 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 100
           ||||||||||||||||||||||||||||||
Db      87 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 116


RESULT 7
AAA12970
ID   AAA12970 standard; DNA; 1126 BP.
XX
AC   AAA12970;
XX
DT   18-JUL-2000  (first entry)
XX
DE   DNA encoding Factor X/Xa, SEQ ID NO:27.
XX
KW   Truncated tissue factor; tTF; human; blood coagulation;
KW   tumour vasculature; bispecific antibody; targetting; cancer;
KW   vascularised tumour; PCR primer; ss.
XX
OS   Homo sapiens.
XX
PN   US6036955-A.
XX
PD   14-MAR-2000.
XX
PF   07-JUN-1995;   95US-0479727.
XX
PR   05-MAR-1992;   92US-0846349.
PR   02-MAR-1994;   94US-0205330.
PR   11-JUL-1994;   94US-0273567.
XX
PA   (TEXA ) UNIV TEXAS SYSTEM.
PA   (SCRI ) SCRIPPS RES INST.
XX
PI   Edgington TS,  Thorpe PE;
XX
DR   WPI; 2000-269871/23.
XX
PT   Kit for inducing coagulation in tumor vasculature, useful for treating
PT   malignant or benign growths, contains ligand, linked to coagulation
PT   agent, that targets tumor marker  -
XX
PS   Example 9; Columns 131-132; 86pp; English.
XX
CC   The invention relates to the induction of blood coagulation specifically
```

CC within tumour vasculature. This is achieved by the use of a bispecific
CC molecule, which comprises a region capable of binding to intratumoral
CC vascular or stromal cells linked to a coagulation factor or to a region
CC capable of binding to a coagulation factor. An example of such a
CC bispecific molecule is a bispecific antibody, where one arm binds a
CC tumour antigen, and the other arm binds a coagulation factor. The
CC expression of certain proteins (tumour antigens) is upregulated in
CC tumour vasculature; such proteins include vascular endothelial growth
CC factor (VEGF) and members of the fibroblast growth factor (FGF) family.
CC An antibody or antibody fragment against VEGF or basic FGF (bFGF) may be
CC incorporated into the bispecific molecule in order to target coagulation
CC to tumour vasculature. The coagulation factor-binding portion of the
CC bispecific molecule may be, for example, directed to tissue factor (TF).
CC A preferred form of TF used in the invention is a truncated form (tTF,
CC AAY81488) which lacks the cytoplasmic and transmembrane domains.
CC Although tTF can associate with Factor VIIa, the tTF/Factor VIIa complex
CC cannot alone initiate the coagulation cascade as the complex has to be
CC associated with a phospholipid surface for coagulation to occur.
CC However, binding of tTF to tumour vasculature via a tumour antigen/tTF
CC bispecific antibody brings tTF into close enough proximity with the cell
CC membrane to enable the initiation of coagulation. Kits for the induction
CC of tumour vasculature-specific coagulation may be used to treat malignant
CC or benign diseases associated with a vascular component, particularly
CC cancers, but also benign growths, prostatic hypertrophy, restenosis,
CC psoriasis, glaucoma, rheumatoid arthritis. Coagulation is induced
CC selectively in the tumour vasculature, minimising side effects. Such kits
CC are likely to be effective against many different types of cancer.
CC Sequences AAA12945- AAA12952, AAA12954-A12963 and AAA12971-A12972
CC represent PCR primers used in exemplifications of the present invention
CC to generate constructs encoding tTF, tTF variants or tTF dimers.
XX
SQ Sequence 1126 BP; 269 A; 341 C; 342 G; 174 T; 0 other;

Query Match 88.4%; Score 88.4; DB 21; Length 1126;
Best Local Similarity 98.9%; Pred. No. 1.9e-18;
Matches 89; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 11 AGTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 70
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 27 ATTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 86

Qy 71 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 100
||||||||||||||||||||||||||||||
Db 87 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 116

RESULT 8
AAZ56120
ID AAZ56120 standard; DNA; 1126 BP.
XX
AC AAZ56120;
XX
DT 27-MAR-2000 (first entry)
XX
DE Vitamin-K-dependent coagulation factor X/Xa coding sequence.
XX
KW Vitamin-K dependent coagulation factor; tumour associated vasculature;
KW carcinoma; benign prostatic hyperplasia; diabetic retinopathy;
KW vascular restenosis; arteriovenous malformation; meningoma; haemangioma;
KW neovascular glaucoma; psoriasis; cytostatic; antidiabetic; vasotropic;
KW ophthalmological; antipsoriatic; Factor X/Xa; ss.
XX
OS Unspecified.
XX
PN US6004555-A.
XX
PD 21-DEC-1999.
XX
PF 07-JUN-1995; 95US-0487427.
XX
PR 05-MAR-1992; 92US-0846349.
PR 02-MAR-1994; 94US-0205330.
PR 11-JUL-1994; 94US-0273567.
XX
PA (SCRI) SCRIPPS RES INST.
PA (TEXA) UNIV TEXAS SYSTEM.
XX
PI Edgington TS, Thorpe PE;
XX
DR WPI; 2000-072047/06.
XX
PT Bispecific binding ligands for promoting blood coagulation in a tumour
PT associated vasculature are useful for treating cancer -
XX
PS Example 9; Column 131-132; 83pp; English.
XX
CC This is the coding sequence for Factor X/Xa, a vitamin-K-dependent
CC coagulation factor. This coagulation factor can be used in the formation
CC of coaguligands. Mutated versions of this sequence can be used in the
CC method for delivering a coagulant to a tumour-associated vasculature
CC using bispecific binding ligands which promote blood coagulation. The
CC binding ligand consists of a binding region that binds to a
CC surface-expressed, surface accessible or surface-localised component of a
CC tumour cell, intratumoural vasculature or tumour associated stroma. The
CC binding region is linked to a coagulating agent which is a coagulation
CC factor (e.g. tissue factor). The second binding region comprises an
CC antibody or an antigen binding region of an antibody. The method is used
CC for delivering an exogenous or an endogenous coagulation factor to
CC tumour-associated vasculature which is benign or malignant. The method
CC can be used to treat cancer by promoting specific blood coagulation in
CC the vasculature of the tumour relative to the vasculature in nontumour
CC sites. Vascularised tumours are usually solid tumours, particularly
CC carcinomas which require a vascular component to provide oxygen and
CC nutrients. The ligands are suitable to treat benign and malignant
CC diseases with a vascular component, including benign prostatic
CC hyperplasia, diabetic retinopathy, vascular restenosis, arteriovenous
CC malformations, meningioma, haemangioma, neovascular glaucoma and
CC psoriasis. The ligands can also be used in standard binding assays in
CC vitro. Bispecific ligands can be designed which are capable of binding to
CC vascular endothelial cells and disease-associated agents such as
CC activated platelets. Certain disease-associated agents are similar in
CC different diseases and in different tumours, making it possible to treat
CC numerous diseases and different types of cancer with one pharmaceutical,
CC therefore an agent does not need to be tailored to each individual
CC disease or specific tumour type.
XX
SQ Sequence 1126 BP; 269 A; 341 C; 342 G; 174 T; 0 other;

Query Match 88.4%; Score 88.4; DB 21; Length 1126;
Best Local Similarity 98.9%; Pred. No. 1.9e-18;
Matches 89; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 11 AGTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 70
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 27 ATTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 86

Qy 71 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 100
||||||||||||||||||||||||||||||
Db 87 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 116

RESULT 9
AAV10462
ID AAV10462 standard; DNA; 1404 BP.
XX
AC AAV10462;
XX
DT 16-JUN-1998 (first entry)
XX
DE Human Factor X protease cDNA.
XX
KW Factor X; factor IX; serine protease activity; catalytic domain; ZAD;
KW zymogen-activating domain; epidermal growth factor-like domain; EGF1;
KW EGF2; regulator; coagulation; fibrinolysis; homeostasis; X-ray structure;
KW detection; drug modelling; restriction protease; ss.
XX

```
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1404
FT                   /*tag= a
FT                   /product= Factor X
FT                   /note= "partial coding sequence"
XX
PN   WO9747737-A1.
XX
PD   18-DEC-1997.
XX
PF   11-JUN-1997;    97WO-EP03027.
XX
PR   06-JUL-1996;    96EP-0110959.
PR   11-JUN-1996;    96EP-0109288.
PR   22-JUN-1996;    96EP-0110109.
XX
PA   (BOEF ) BOEHRINGER MANNHEIM GMBH.
XX
PI   Hopfner K,  Kopetzki E;
XX
DR   WPI; 1998-052304/05.
DR   P-PSDB; AAW40283.
XX
PT   Non-glycosylated, truncated forms of factor IX family protein with
PT   serine protease activity - used to screen for specific modulators
PT   and to assay factor IXa
XX
PS   Disclosure; Fig 3; 49pp; German.
XX
CC   This sequence encodes a human factor X protease. This protein is used
CC   in the construction of a novel non-glycosylated protein and truncated
CC   and zymogen forms of this protein, which have serine protease activity.
CC   The protein is composed of various domains from a factor IX family
CC   protein, namely a catalytic domain (CD) N-terminally bound to a
CC   zymogen-activating domain (ZAD), N-terminally bound to an EGF1 and/or
CC   EGF2 domain (EGF = epidermal growth factor-like domain). Such proteins
CC   are used to identify activators/inhibitors of factor IX family proteins
CC   (potentially useful as regulators of coagulation, fibrinolysis and
CC   homeostasis). The protein in zymogen form is also useful in assays for
CC   detecting factor IXa activity in aqueous solution (specifically in body
CC   fluids). The protein can be used to produce co-crystals with protease
CC   variants or inhibitors for X-ray structural analysis and drug modelling
CC   and as restriction proteases in biotechnology. These truncated proteins
CC   have the same specificity as factor IX family proteases and can be
CC   produced in prokaryotes in a form that allows production of active enzyme
CC   by conversion to native form and enzymatic cleavage.

     Sequence 1404 BP; 356 A; 404 C; 423 G; 221 T; 0 other;

  Query Match             88.4%;  Score 88.4;  DB 19;  Length 1404;
  Best Local Similarity   98.9%;  Pred. No. 2e-18;
  Matches   89;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      11 AGTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 70
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     315 ATTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 374

Qy      71 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 100
           ||||||||||||||||||||||||||||||
Db     375 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 404


RESULT 10
AAV56776
ID   AAV56776 standard; DNA; 1467 BP.
XX
AC   AAV56776;
XX
DT   27-NOV-1998  (first entry)
XX
DE   Human Factor X genomic DNA.
XX
KW   Factor X; analogue; activation cleavage site; protease; bleeding; human;
KW   Factor IX; Factor VII; Factor VIII; haemophilia; gene therapy; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1467
FT                   /*tag= a
FT                   /product= "Factor X"
FT   sig_peptide     1..120
FT                   /*tag= b
FT   mat_peptide     121..1464
FT                   /*tag= c
XX
PN   WO9838317-A1.
XX
PD   03-SEP-1998.
XX
PF   27-FEB-1998;    98WO-AT00045.
XX
PR   27-FEB-1997;    97AT-0000335.
XX
PA   (IMMO ) IMMUNO AG.
XX
PI   Dorner F,  Eibl J,  Fisch A,  Himmelspach M,  Schlokat U;
XX
DR   WPI; 1998-481211/41.
DR   P-PSDB; AAW76216.
XX
PT   New factor X analogues with processing site for protease not active
PT   on natural protein - and related DNA, is very stable and can be
PT   activated in vitro or in vivo without using animal protease(s),
PT   particularly for treating disorders of blood coagulation
XX
PS   Claim 3; Fig 1; 86pp; German.
XX
CC   This sequence encodes the human Factor X protein which is used in a
CC   method resulting in the production of novel human Factor X (F10)
CC   analogues. Such analogues have in the region of the natural F10a
CC   activation cleavage site, a modification that creates a processing site
CC   for a protease that does not naturally cleave F10 in this region. The
CC   proteins are used to generate, in vivo or in vitro, F10a analogues that
CC   can be used to control bleeding and for treating defects of factors IX,
CC   VII or VIII, e.g.in haemophiliacs who have developed antibodies to
CC   factors VIII and/or IX. The encoding nucleic acid can be used in gene
CC   therapy of the same conditions. The analogues have high stability and can
CC   be activated without use of animal enzymes such as trypsin. Only
CC   activation is affected, their activity is the same as the natural factor.
CC   The analogues can be isolated as a pure single-chain pro-protein (not
CC   usually possible because of rapid processing of the native precursor) and
CC   this converted to two-chain form by subsequent activation. Activated
CC   analogues have good stability and structural integrity and are
CC   practically free of inactive intermediates and autoproteolytic
CC   decomposition products.
XX
SQ   Sequence 1467 BP; 363 A; 424 C; 444 G; 236 T; 0 other;

  Query Match             88.4%;  Score 88.4;  DB 19;  Length 1467;
  Best Local Similarity   98.9%;  Pred. No. 2e-18;
  Matches   89;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      11 AGTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 70
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     369 ATTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 428

Qy      71 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 100
           ||||||||||||||||||||||||||||||
Db     429 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 458


RESULT 11
AAV56821
```

```
ID   AAV56821 standard; DNA; 1467 BP.
XX
AC   AAV56821;
XX
DT   27-NOV-1998  (first entry)
XX
DE   Human Factor X genomic DNA.
XX
KW   Factor X; analogue; activation cleavage site; protease; bleeding; human;
KW   Factor IX; Factor VII; Factor VIII; haemophilia; gene therapy; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1467
FT                   /*tag=  a
FT   sig_peptide     1..120
FT                   /*tag=  b
FT   mat_peptide     121..1464
FT                   /*tag=  c
FT                   /product= "Factor X"
XX
PN   WO9838318-A1.
XX
PD   03-SEP-1998.
XX
PF   27-FEB-1998;   98WO-AT00046.
XX
PR   27-FEB-1997;   97AT-0000336.
XX
PA   (IMMO ) IMMUNO AG.
XX
PI   Dorner F,  Eibl J,  Falkner F,  Himmelspach M,  Pfleiderer M;
PI   Schlokat U;
XX
DR   WPI; 1998-481212/41.
DR   P-PSDB; AAW76218.
XX
PT   New factor 10 deletion mutants lacking the natural protease
PT   processing site - but having a non-natural site inserted, and
PT   related DNA, particularly for in vitro activation to products used
PT   to treat blood coagulation disorders
XX
PS   Claim 3; Fig 1; 82pp; German.
XX
CC   This sequence encodes the human Factor X protein which is used in a
CC   method resulting in the production of novel human Factor X (F10)
CC   analogues. Such analogues have in the region of the natural F10a
CC   activation cleavage site, a modification that creates a processing site
CC   for a protease that does not naturally cleave F10 in this region. The
CC   proteins are used to generate, in vivo or in vitro, F10a analogues that
CC   can be used to control bleeding and for treating defects of factors IX,
CC   VII or VIII, e.g.in haemophiliacs who have developed antibodies to
CC   factors VIII and/or IX. The encoding nucleic acid can be used in gene
CC   therapy of the same conditions. The analogues have high stability and can
CC   be activated without use of animal enzymes such as trypsin. Only
CC   activation is affected, their activity is the same as the natural factor.
CC   The analogues can be isolated as a pure single-chain pro-protein (not
CC   usually possible because of rapid processing of the native precursor) and
CC   this converted to two-chain form by subsequent activation. Activated
CC   analogues have good stability and structural integrity and are
CC   practically free of inactive intermediates and autoproteolytic
CC   decomposition products.
XX
SQ   Sequence 1467 BP; 363 A; 424 C; 444 G; 236 T; 0 other;

  Query Match            88.4%;  Score 88.4;  DB 19;  Length 1467;
  Best Local Similarity  98.9%;  Pred. No. 2e-18;
  Matches   89;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      11 AGTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 70
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     369 ATTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 428

Qy      71 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 100
           ||||||||||||||||||||||||||||||
Db     429 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 458


RESULT 12
AAF59409
ID   AAF59409 standard; cDNA; 1467 BP.
XX
AC   AAF59409;
XX
DT   02-MAY-2001  (first entry)
XX
DE   Human factor X nucleotide sequence SEQ ID NO:1.
XX
KW   Human; factor X; mutant; haemostatic; gene therapy; haemophilia;
KW   blood coagulation disorder; haemophiliac; ss.
XX
OS   Homo sapiens.
XX
PN   WO200110896-A2.
XX
PD   15-FEB-2001.
XX
PF   07-AUG-2000; 2000WO-EP07631.
XX
PR   10-AUG-1999;   99AT-0001377.
XX
PA   (BAXT ) BAXTER AG.
XX
PI   Himmelspach M,  Schlokat U;
XX
DR   WPI; 2001-191516/19.
DR   P-PSDB; AAB70411.
XX
PT   Novel factor X analog useful for producing drug which is useful for
PT   treatment of blood coagulation disorders, such as hemophilia, contains
PT   modification between amino acids Glu226 and Ile235  -
XX
PS   Disclosure; Fig 1; 50pp; English.
XX
CC   The present invention describes a factor X analogue (I) which contains
CC   a modification between Glu226 and Ile235, relative to the 488 residue
CC   amino acid sequence given in AAB70411. (I) has haemostatic activity and
CC   can be used in gene therapy. (I) encoding polynucleotide (II) can be
CC   used to produce a drug, which is useful for treatment of patients with
CC   blood coagulation disorders, such as patients suffering from haemophilia,
CC   or haemophiliacs with inhibitory antibodies. Preparations containing a
CC   polypeptide with factor X/Xa activity are more readily activated by
CC   factor X1a or its derivative, which has high stability, without having
CC   to use one of the proteases used in prior art to activate the natural
CC   factor X, particularly one of animal origins, such as Russell's viper
CC   venom (RVV) or trypsin. The present sequence encodes human factor X,
CC   which is given in the exemplification of the present invention.
XX
SQ   Sequence 1467 BP; 363 A; 424 C; 444 G; 236 T; 0 other;

  Query Match            88.4%;  Score 88.4;  DB 22;  Length 1467;
  Best Local Similarity  98.9%;  Pred. No. 2e-18;
  Matches   89;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      11 AGTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 70
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     369 ATTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 428

Qy      71 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 100
           ||||||||||||||||||||||||||||||
Db     429 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 458


RESULT 13
AAA54031
```

```
ID   AAA54031 standard; DNA; 1507 BP.
XX
AC   AAA54031;
XX
DT   08-FEB-2001  (first entry)
XX
DE   Human factor X coding sequence.
XX
KW   Vitamin K dependent protein; VKDP; gamma-carboxylation; chimeric
KW   protein; fusion protein; coagulation factor; Factor X; Factor VII;
KW   Protein S; Factor IX; Protein C; prothrombin; blood clotting;
KW   haemophilia; human; ds.
XX
OS   Homo sapiens.
XX
PN   WO200054787-A1.
XX
PD   21-SEP-2000.
XX
PF   16-MAR-2000; 2000WO-US06934.
XX
PR   16-MAR-1999;   99US-0124609.
XX
PA   (CHIL-) CHILDRENS HOSPITAL PHILADELPHIA.
PA   (UYNC-) UNIV NORTH CAROLINA.
XX
PI   High KA,  Camire RM,  Larson PJ,  Stafford DW;
XX
DR   WPI; 2000-638152/61.
XX
PT   Chimeric DNA for optimizing gamma carboxylation of vitamin K-dependent
PT   protein useful for treating diseases associated with the protein,
PT   comprises sequence encoding propeptide fused to sequence encoding the
PT   protein
XX
PS   Disclosure; Fig 6a; 60pp; English.
XX
CC   Efficient processing and release of mature two-chain factor X into
CC   the circulation requires: removal of the signal sequence; formation
CC   of disulfide bonds; modification of amino terminal glutamic acid
CC   residues, to gamma-carboxyglutamic acid; modification of one
CC   aspartic acid in the first epidermal growth factor (EGF) domain to
CC   Beta-hydroxyaspartic acid; addition of N- and O-linked
CC   oligosaccharides to the activation peptide; removal of an internal
CC   tripeptide to yield two-chain factor X and removal of the
CC   propeptide just prior to secretion. While some of these modifications
CC   do not appear essential for factor X function the removal of the
CC   signal sequence, propeptide, internal tripeptide and full
CC   gamma-carboxylation are all steps which are important requisites for
CC   the production of biologically active factor X/FXa. Isolated chimeric
CC   polynucleotides are described which encode a propeptide fused to a
CC   nucleic acid sequence encoding a vitamin K-dependent protein (VKDP).
CC   The fusion proteins encoded are vitamin K-dependent protein
CC   gamma-carboxylation enhancers and are useful for optimising the
CC   gamma-carboxylation of a VKDP to produce a fully gamma-carboxylated
CC   VKDP. The fusion proteins and recombinant cells expressing them are
CC   useful for alleviating a VKDP associated disease. The fusion
CC   constructs result in the production of fully gamma-carboxylated
CC   mature VKDPs  which are biologically active. The invention
CC   encompasses all combinations of propeptide sequences (modified or
CC   not) and VKDP's. This sequence encodes the signal, propeptide and
CC   mature protein sequence of human Factor X.
XX
SQ   Sequence 1507 BP; 394 A; 429 C; 446 G; 238 T; 0 other;

  Query Match             88.4%;  Score 88.4;  DB 21;  Length 1507;
  Best Local Similarity   98.9%;  Pred. No. 2e-18;
  Matches   89;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      11 AGTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 70
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     369 ATTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 428

Qy      71 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 100
           ||||||||||||||||||||||||||||||
Db     429 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 458


RESULT 14
AAH57469
ID   AAH57469 standard; cDNA; 1887 BP.
XX
AC   AAH57469;
XX
DT   10-SEP-2001  (first entry)
XX
DE   Human liver cell specific cDNA sequence SEQ ID NO:309.
XX
KW   Human; tissue specific; diagnosis; brain; heart; skeletal muscle;
KW   lung; liver; uterus; ovary; stomach; intestine; kidney; pancreas; ss;
KW   metabolic disease; developmental disease; cytostatic; immunomodulatory;
KW   neuroprotective; gene therapy; cancer; immunopathology; neuropathology.
XX
OS   Homo sapiens.
XX
PN   WO200132927-A2.
XX
PD   10-MAY-2001.
XX
PF   02-NOV-2000; 2000WO-US30396.
XX
PR   04-NOV-1999;   99US-0163508.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Sornasse T,  Seilhamer JJ,  Watson GA;
XX
DR   WPI; 2001-291057/30.
XX
PT   New cell and tissue specific polynucleotides useful for diagnosis,
PT   prognosis or monitoring of treatments for disorders where the gene is
PT   associated with a cancer, immunopathology or neuropathology -
XX
PS   Claim 1; Page 233; 327pp; English.
XX
CC   AAH57161 to AAH57576 represent cell and tissue specific polynucleotide
CC   sequences (I). (I) can have cytostatic, immunomodulatory and
CC   neuroprotective activities, and can be used in gene therapy. (I) and
CC   proteins (II) encoded by then are used in high throughput screening
CC   assays to select DNA molecules, RNA molecules, peptide nucleic acids,
CC   mimetics, peptides, proteins, agonists, antagonists, antibodies or
CC   their fragments, immunoglobulins, inhibitors, drug compounds and
CC   pharmaceutical agents. Expression of (I) in a sample indicates the
CC   differentiation of embryonic stem cells into a tissue selected from
CC   brain, heart, kidney, liver, lung, skeletal muscle or pancreatic
CC   tissues. (I) and (II) are used to produce an expression profile that
CC   defines a metabolic or developmental process, treatment, condition,
CC   disease or disorder. The gene profile can be used for diagnosis,
CC   prognosis or monitoring of treatments and for investigating a
CC   predisposition to a disorder where the gene is associated with a
CC   cancer, immunopathology or neuropathology.
XX
SQ   Sequence 1887 BP; 467 A; 549 C; 544 G; 327 T; 0 other;

  Query Match             88.4%;  Score 88.4;  DB 22;  Length 1887;
  Best Local Similarity   98.9%;  Pred. No. 2.1e-18;
  Matches   89;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      11 AGTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 70
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     424 ATTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 483

Qy      71 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 100
           ||||||||||||||||||||||||||||||
Db     484 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 513
```

```
RESULT 15
AAH57261
ID   AAH57261 standard; cDNA; 300 BP.
XX
AC   AAH57261;
XX
DT   10-SEP-2001  (first entry)
XX
DE   Human liver specific cDNA sequence SEQ ID NO:101.
XX
KW   Human; tissue specific; diagnosis; brain; heart; skeletal muscle;
KW   lung; liver; uterus; ovary; stomach; intestine; kidney; pancreas; ss;
KW   metabolic disease; developmental disease; cytostatic; immunomodulatory;
KW   neuroprotective; gene therapy; cancer; immunopathology; neuropathology.
XX
OS   Homo sapiens.
XX
PN   WO200132927-A2.
XX
PD   10-MAY-2001.
XX
PF   02-NOV-2000; 2000WO-US30396.
XX
PR   04-NOV-1999;   99US-0163508.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Sornasse T,  Seilhamer JJ,  Watson GA;
XX
DR   WPI; 2001-291057/30.
XX
PT   New cell and tissue specific polynucleotides useful for diagnosis,
PT   prognosis or monitoring of treatments for disorders where the gene is
PT   associated with a cancer, immunopathology or neuropathology -
XX
PS   Claim 1; Page 107; 327pp; English.
XX
CC   AAH57161 to AAH57576 represent cell and tissue specific polynucleotide
CC   sequences (I). (I) can have cytostatic, immunomodulatory and
CC   neuroprotective activities, and can be used in gene therapy. (I) and
CC   proteins (II) encoded by then are used in high throughput screening
CC   assays to select DNA molecules, RNA molecules, peptide nucleic acids,
CC   mimetics, peptides, proteins, agonists, antagonists, antibodies or
CC   their fragments, immunoglobulins, inhibitors, drug compounds and
CC   pharmaceutical agents. Expression of (I) in a sample indicates the
CC   differentiation of embryonic stem cells into a tissue selected from
CC   brain, heart, kidney, liver, lung, skeletal muscle or pancreatic
CC   tissues. (I) and (II) are used to produce an expression profile that
CC   defines a metabolic or developmental process, treatment, condition,
CC   disease or disorder. The gene profile can be used for diagnosis,
CC   prognosis or monitoring of treatments and for investigating a
CC   predisposition to a disorder where the gene is associated with a
CC   cancer, immunopathology or neuropathology.
XX
SQ   Sequence 300 BP; 73 A; 81 C; 88 G; 44 T; 14 other;

  Query Match             85.4%;  Score 85.4;  DB 22;  Length 300;
  Best Local Similarity   95.6%;  Pred. No. 1.2e-17;
  Matches   86;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      11 AGTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 70
           | | | ||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db      16 ATTNANACGGAAGCTCTGNAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 75

Qy      71 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 100
           ||||||||||||||||||||||||||||||
Db      76 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 105


Search completed: January 15, 2003, 19:50:15
Job time : 154 secs
```

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: January 15, 2003, 17:56:05 ; Search time 1334 Seconds
(without alignments)
2181.620 Million cell updates/sec

Title: L00394_COPY_1_100
Perfect score: 100
Sequence: 1 CTCCTTTGGCAGTCACACGG.....GTGGTGTGCTCCTGCGCCCG 100

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	100	100.0	156	9	HUMFX5	L00394 Human facto
2	93.8	93.8	29488	9	AF503510	AF503510 Homo sapi
3	93.8	93.8	132933	9	AL137002	AL137002 Human DNA
4	92	92.0	350	9	AB005892	AB005892 Homo sapi
5	88.4	88.4	1126	6	AR095306	AR095306 Sequence
6	88.4	88.4	1126	6	AR103990	AR103990 Sequence
7	88.4	88.4	1126	9	HUMFX	K01886 Human facto
8	88.4	88.4	1404	6	A93124	A93124 Sequence 15
9	88.4	88.4	1414	9	HUMCFX	M22613 Human blood
10	88.4	88.4	1443	9	HUMFXM	K03194 Human facto
11	88.4	88.4	1467	6	A86859	A86859 Sequence 43
12	88.4	88.4	1467	6	A86886	A86886 Sequence 26
13	88.4	88.4	1467	6	AX082959	AX082959 Sequence
14	88.4	88.4	1500	6	AR024194	AR024194 Sequence
15	88.4	88.4	1507	9	HUMFACX	M57285 Human coagu
16	72.2	72.2	49581	2	AC131472	AC131472 Rattus no
17	69.4	69.4	20130	10	AF211347	AF211347 Mus muscu
18	69.4	69.4	299880	2	AC127308	AC127308 Mus muscu
19	68.6	68.6	1500	4	AF003200	AF003200 Oryctolag
20	66.6	66.6	1497	10	RNFXRAT	X79807 R.norvegicu
21	65.4	65.4	1486	10	AF087644	AF087644 Mus muscu
22	65.4	65.4	1491	10	MMCOAGULX	AJ222677 Mus muscu
23	65.4	65.4	1925	10	BC003877	BC003877 Mus muscu
24	52.4	52.4	1537	4	BTCFXI	X00673 Bovine mRNA
25	52.4	52.4	1554	6	A73583	A73583 Sequence 1
26	52.4	52.4	1554	6	AR001423	AR001423 Sequence
27	45.4	45.4	1599	4	AF275654	AF275654 Ornithorh
28	40.2	40.2	20947	5	AF519546	AF519546 Danio rer
29	32.4	32.4	2246	5	CHKVAP	D00844 Gallus gall
30	31.8	31.8	132171	3	AC008370	AC008370 Drosophil
31	31.8	31.8	154370	2	AC108106	AC108106 Homo sapi
32	31.8	31.8	157559	2	AC020344	AC020344 Drosophil
33	31.8	31.8	226833	3	AE003838	AE003838 Drosophil
34	30.8	30.8	1558	4	OCU49933	U49933 Oryctolagus
35	30.8	30.8	2461	6	I23785	I23785 Sequence 1
36	30.8	30.8	2461	9	HUMGAS	L13720 Homo sapien
c 37	30.6	30.6	126721	2	AC118793	AC118793 Rattus no
c 38	30.6	30.6	159743	2	AC126937	AC126937 Mus muscu
c 39	30.2	30.2	148198	2	AC024232	AC024232 Homo sapi
40	30.2	30.2	218138	2	AC122193	AC122193 Mus muscu
41	30	30.0	96642	8	AC002291	AC002291 Arabidops
c 42	30	30.0	203200	9	AC008744	AC008744 Homo sapi
43	29.6	29.6	6846	9	AB011540	AB011540 Homo sapi
c 44	29.6	29.6	113228	9	AL513015	AL513015 Human DNA
45	29.6	29.6	114191	9	AL451007	AL451007 Human DNA

ALIGNMENTS

RESULT 1
HUMFX5
LOCUS HUMFX5 156 bp DNA linear PRI 09-NOV-1994
DEFINITION Human factor X (blood coagulation factor) gene, exon 5.
ACCESSION L00394 M14327
VERSION L00394.1 GI:182826
KEYWORDS Stuart factor; blood coagulation factor; factor X; glycoprotein; serine protease.
SEGMENT 5 of 8
SOURCE Homo sapiens (tissue library: of Lawn et al., and Yoshitake et al.) DNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 156)

```
  AUTHORS   Leytus,S.P., Foster,D.C., Kurachi,K. and Davie,E.W.
  TITLE     Gene for human factor X: a blood coagulation factor whose gene
            organization is essentially identical with that of factor IX and
            protein C
  JOURNAL   Biochemistry 25 (18), 5098-5102 (1986)
  MEDLINE   87026600
   PUBMED   3768336
FEATURES             Location/Qualifiers
     source          1. .156
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /map="13q34"
                     /tissue_lib="of Lawn et al., and Yoshitake et al."
     intron          <1. .12
                     /gene="F10"
                     /note="FX intron D"
     exon            13. .144
                     /gene="F10"
                     /note="G00-119-890"
                     /number=5
     intron          145. .>156
                     /gene="F10"
                     /note="FX intron E"
BASE COUNT       32 a     49 c     48 g     27 t
ORIGIN      About 1.8 kb after segment 4; chromosome 13q34.

  Query Match             100.0%;  Score 100;  DB 9;  Length 156;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-19;
  Matches  100;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CTCCTTTGGCAGTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCT 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CTCCTTTGGCAGTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCT 60

Qy      61 GCCACGAGGAACAGAACTCTGTGGTGTGCTCCTGCGCCCG 100
           ||||||||||||||||||||||||||||||||||||||||
Db      61 GCCACGAGGAACAGAACTCTGTGGTGTGCTCCTGCGCCCG 100


RESULT 2
AF503510
LOCUS       AF503510               29488 bp    DNA     linear   PRI 22-MAY-2002
DEFINITION  Homo sapiens coagulation factor X (F10) gene, complete cds.
ACCESSION   AF503510
VERSION     AF503510.1  GI:20336662
KEYWORDS    .
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 29488)
  AUTHORS   Rieder,M.J., Armel,T.Z., Carrington,D.P., Chung,M.-W., Lee,K.L.,
            Ozuna,M., Poel,C.L., Toth,E.J., Yi,Q. and Nickerson,D.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-APR-2002) Genome Sciences, University of Washington,
            1705 NE Pacific, Seattle, WA 98195, USA
COMMENT     To cite this work please use:  SeattleSNPs. NHLBI HL66682 Program
            for Genomic Applications, UW-FHCRC, Seattle, WA (URL:
            http://pga.gs.washington.edu).
FEATURES             Location/Qualifiers
     source          1. .29488
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     variation       152
                     /frequency="0.04"
                     /replace="c"
     variation       167
                     /frequency="0.01"
                     /replace="t"
     variation       370
                     /frequency="0.01"
                     /replace="a"
     variation       577
                     /frequency="0.33"
                     /replace="t"
     variation       671
                     /frequency="0.01"
                     /replace="t"
     variation       713
                     /frequency="0.12"
                     /replace="c"
     variation       851
                     /frequency="0.01"
                     /replace="g"
     variation       866
                     /frequency="0.02"
                     /replace="a"
     variation       880
                     /frequency="0.05"
                     /replace="c"
     variation       1116
                     /frequency="0.79"
                     /replace=""
     variation       1169
                     /frequency="0.02"
                     /replace="a"
     repeat_region   1189. .1395
                     /rpt_family="MIR"
                     /rpt_type=dispersed
     variation       1319. .1324
                     /frequency="0.57"
                     /replace=""
     variation       1449
                     /frequency="0.48"
                     /replace="c"
     variation       1451
                     /frequency="0.33"
                     /replace="a"
     variation       1612
                     /frequency="0.01"
                     /replace="a"
     variation       1632
                     /frequency="0.01"
                     /replace="t"
     gene            <1647. .28346
                     /gene="F10"
     mRNA            join(<1647. .1741,8269. .8429,17274. .17298,18174. .18287,
                     19736. .19867,22668. .22912,26198. .26315,27735. .28346)
                     /gene="F10"
                     /product="coagulation factor X"
     CDS             join(1672. .1741,8269. .8429,17274. .17298,18174. .18287,
                     19736. .1986?,22668. .22912,26198. .26315,27735. .28336)
                     /gene="F10"
                     /codon_start=1
                     /product="coagulation factor X"
                     /protein_id="AAM19347.1"
                     /db_xref="GI:20336663"
                     /translation="MGRPLHLVLLSASLAGLLLLGESLFIRREQANNILARVTRANSF
                     LEEMKKGHLERECMEETCSYEEAREVFEDSDKTNEFWNKYKDGDQCETSPCQNQGKCK
                     DGLGEYTCTCLEGFEGKNCELFTRKLCSLDNGDCDQFCHEEQNSVVCSCARGYTLADN
                     GKACIPTGPYPCGKQTLERRKRSVAQATSSSGEAPDSITWKPYDAADLDPTENPFDLL
                     DFNQTQPERGDNNLTRIVGGQECKDGECPWQALLINEENEGFCGGTILSEFYILTAAH
                     CLYQAKRFKVRVGDRNTEQEEGGEAVHEVEVVIKHNRFTKETYDFDIAVLRLKTPITF
                     RMNVAPACLPERDWAESTLMTQKTGIVSGFGRTHEKGRQSTRLKMLEVPYVDRNSCKL
                     SSSFIITQNMFCAGYDTKQEDACQGDSGGPHVTRFKDTYFVTGIVSWGEGCARKGKYG
                     IYTKVTAFLKWIDRSMKTRGLPKAKSHAPEVITSSPLK"
     variation       1731
                     /gene="F10"
                     /frequency="0.01"
                     /replace="t"
     variation       2011
                     /gene="F10"
                     /frequency="0.22"
                     /replace="g"
     variation       2456
                     /gene="F10"
```

```
                /frequency="0.04"
                /replace="a"
variation       2527
                /gene="F10"
                /frequency="0.12"
                /replace="c"
variation       2757
                /gene="F10"
                /frequency="0.04"
                /replace="g"
misc_feature    2860. .3410
                /gene="F10"
                /note="Region not scanned for variation"
repeat_region   3544. .3975
                /rpt_family="MaLR"
                /rpt_type=dispersed
variation       3715
                /gene="F10"
                /frequency="0.09"
                /replace="t"
variation       3740
                /gene="F10"
                /frequency="0.39"
                /replace="g"
variation       3977
                /gene="F10"
                /frequency="0.01"
                /replace="c"
variation       4100
                /gene="F10"
                /frequency="0.01"
                /replace="t"
variation       4544
                /gene="F10"
                /frequency="0.19"
                /replace="c"
repeat_region   4567. .4967
                /rpt_family="L2"
                /rpt_type=dispersed
variation       5128
                /gene="F10"
                /frequency="0.01"
                /replace="g"
variation       5300
                /gene="F10"
                /frequency="0.02"
                /replace="c"
variation       5413
                /gene="F10"
                /frequency="0.13"
                /replace="c"
variation       5521
                /gene="F10"
                /frequency="0.11"
                /replace="a"
misc_feature    5540. .5950
                /gene="F10"
                /note="Region not scanned for variation"
variation       5997
                /gene="F10"
                /frequency="0.01"
                /replace="c"
variation       6098
                /gene="F10"
                /frequency="0.03"
                /replace="a"
variation       6110
                /gene="F10"
                /frequency="0.01"
                /replace="c"
variation       6445
                /gene="F10"
                /frequency="0.06"
                /replace="c"
variation       6638. .6641
                /gene="F10"
                /frequency="0.99"
                /replace=""
repeat_region   6943. .7395
                /rpt_family="ERV1"
                /rpt_type=dispersed
repeat_region   7396. .7727
                /rpt_family="Alu"
                /rpt_type=dispersed
variation       7437
                /gene="F10"
                /frequency="0.01"
                /replace="a"
repeat_region   7728. .8145
                /rpt_family="ERV1"
                /rpt_type=dispersed
variation       7742
                /gene="F10"
                /frequency="0.01"
                /replace="t"

  Query Match             93.8%;  Score 93.8;  DB 9;  Length 29488;
  Best Local Similarity   97.9%;  Pred. No. 7.8e-18;
  Matches   95;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 CTTTGGCAGTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCC 63
           | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   19727 CCTTTGCAGTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCC 19786

Qy      64 ACGAGGAACAGAACTCTGTGGTGTGCTCCTGCGCCCG 100
           |||||||||||||||||||||||||||||||||||||
Db   19787 ACGAGGAACAGAACTCTGTGGTGTGCTCCTGCGCCCG 19823

RESULT 3
AL137002
LOCUS       AL137002              132933 bp    DNA     linear   PRI 04-MAY-2001
DEFINITION  Human DNA sequence from clone RP11-98F14 on chromosome
            13q22.1-31.1, complete sequence.
ACCESSION   AL137002
VERSION     AL137002.19  GI:13990998
KEYWORDS    HTG.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 132933)
  AUTHORS   Bates,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-MAY-2001) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT     On May 8, 2001 this sequence version replaced gi:13274222.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest. The following
            abbreviations are used to associate primary accession numbers given
            in the feature table with their source databases: Em:, EMBL; Sw:,
            SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP
            database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
```

```
            was generated from part of bacterial clone contigs of human
            chromosome 13, constructed by the Sanger Centre Chromosome 13
            Mapping Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr13
            RP11-98F14 is from the library RPCI-11.1 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6
            IMPORTANT: This sequence is not the entire insert of clone
            RP11-98F14 It may be shorter because we sequence overlapping
            sections only once, except for a 100 base overlap.
            The true left end of clone RP11-98F14 is at 1 in this sequence. The
            true left end of clone RP11-391H12 is at 132834 in this sequence.
            The true right end of clone RP11-265C7 is at 123923 in this
            sequence.
FEATURES                Location/Qualifiers
     source             1. .132933
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /chromosome="13"
                        /map="q22.1-31.1"
                        /clone="RP11-98F14"
                        /clone_lib="RPCI-11.1"
     repeat_region      41. .277
                        /note="AluJo repeat: matches 1. .229 of consensus"
     repeat_region      377. .472
                        /note="3 copies 32 mer 93% conserved"
     repeat_region      901. .1207
                        /note="AluSq repeat: matches 1. .307 of consensus"
     misc_feature       1483. .1912
                        /note="CpG island"
                        /evidence=not_experimental
     repeat_region      1999. .2164
                        /note="MER58A repeat: matches 19. .167 of consensus"
     repeat_region      2703. .2861
                        /note="L1M3e repeat: matches -118. .34 of consensus"
     repeat_region      4371. .4978
                        /note="38 copies 16 mer 78% conserved"
     repeat_region      4374. .4981
                        /note="19 copies 32 mer 87% conserved"
     repeat_region      6373. .6448
                        /note="2 copies 38 mer 93% conserved"
     repeat_region      6411. .6563
                        /note="3 copies 51 mer 93% conserved"
     repeat_region      6565. .6669
                        /note="3 copies 35 mer 85% conserved"
     repeat_region      6612. .6687
                        /note="2 copies 38 mer 98% conserved"
     repeat_region      6672. .6773
                        /note="2 copies 51 mer 93% conserved"
     repeat_region      8852. .8923
                        /note="4 copies 18 mer 81% conserved"
     repeat_region      8854. .8988
                        /note="15 copies 9 mer cccgaaccc 63% conserved"
     repeat_region      8860. .8943
                        /note="3 copies 28 mer 83% conserved"
     repeat_region      8920. .9059
                        /note="4 copies 35 mer 75% conserved"
     repeat_region      8927. .9056
                        /note="5 copies 26 mer 69% conserved"
     repeat_region      8935. .9002
                        /note="4 copies 17 mer 95% conserved"
     repeat_region      8984. .9037
                        /note="3 copies 18 mer 90% conserved"
     repeat_region      9427. .9530
                        /note="4 copies 26 mer 92% conserved"
     repeat_region      10204. .10602
                        /note="7 copies 57 mer 87% conserved"
     repeat_region      11050. .11091
                        /note="2 copies 21 mer 100% conserved"
     repeat_region      11469. .11823
                        /note="L1MC1 repeat: matches 5862. .6231 of consensus"
     repeat_region      13157. .13200
                        /note="LTR5 repeat: matches 690. .733 of consensus"
     repeat_region      13806. .14155
                        /note="5 copies 70 mer 82% conserved"
     misc_feature       15491. .15973
                        /note="CpG island"
                        /evidence=not_experimental
     repeat_region      16278. .16371
                        /note="2 copies 47 mer 87% conserved"
     repeat_region      16956. .17159
                        /note="4 copies 51 mer 95% conserved"
     repeat_region      19411. .19587
                        /note="3 copies 59 mer 79% conserved"
     repeat_region      20202. .20342
                        /note="3 copies 47 mer 96% conserved"
     repeat_region      20473. .20564
                        /note="L2 repeat: matches 2602. .2693 of consensus"
     misc_feature       20816. .21894
                        /note="CpG island"
                        /evidence=not_experimental
     repeat_region      24570. .24680
                        /note="L1M4 repeat: matches 3584. .3686 of consensus"
     repeat_region      24681. .24999
                        /note="AluJo repeat: matches 1. .311 of consensus"
     repeat_region      25000. .25306
                        /note="L1M4 repeat: matches 3686. .4029 of consensus"
     repeat_region      25311. .25620
                        /note="AluY repeat: matches 1. .311 of consensus"
     misc_feature       25345. .25989
                        /note="CpG island"
                        /evidence=not_experimental
     repeat_region      25755. .26053
                        /note="AluSq repeat: matches 1. .304 of consensus"
     repeat_region      27585. .27688
                        /note="4 copies 26 mer 74% conserved"
     repeat_region      27733. .27981
                        /note="AluJo repeat: matches 9. .282 of consensus"
     repeat_region      29431. .29498
                        /note="MER81 repeat: matches 4. .72 of consensus"
     repeat_region      29553. .29870
                        /note="AluSx repeat: matches 1. .309 of consensus"
     repeat_region      30907. .31396
                        /note="7 copies 70 mer 64% conserved"
     repeat_region      30908. .31432
                        /note="15 copies 35 mer 61% conserved"
     repeat_region      30911. .31412
                        /note="251 copies 2 mer gg 61% conserved"
     repeat_region      30938. .31413
                        /note="17 copies 28 mer 61% conserved"
     repeat_region      30973. .31420
                        /note="14 copies 32 mer 61% conserved"
     repeat_region      30987. .31385
                        /note="7 copies 57 mer 64% conserved"
     misc_feature       31000. .31080
                        /note="Sequence from uni-directional dGTP big dye
                        terminator reads only."
     repeat_region      31043. .31413
                        /note="7 copies 53 mer 63% conserved"
     repeat_region      31078. .31415
                        /note="13 copies 26 mer 62% conserved"
     repeat_region      31136. .31417
                        /note="6 copies 47 mer 66% conserved"
     repeat_region      31152. .31421
                        /note="9 copies 30 mer 61% conserved"
     misc_feature       31153. .31267
                        /note="Single clone region. Sequence from reads from a
                        short insert library derived from a single pUC clone.
                        Restriction digest data confirm the assembly."
     repeat_region      31175. .31426
                        /note="12 copies 21 mer 61% conserved"
     repeat_region      31268. .31411
                        /note="9 copies 16 mer 64% conserved"
     misc_feature       31268
                        /note="Tandem repeat. Forced join. Gap size estimated to
```

```
                be approximately 350bp by restriction digest data."
     repeat_region   31882. .32952
                     /note="21 copies 51 mer 89% conserved"
     repeat_region   31884. .32937
                     /note="62 copies 17 mer 89% conserved"
     misc_feature    31896. .35863
                     /note="CpG island"
                     /evidence=not_experimental
     repeat_region   31982. .33941
                     /note="28 copies 70 mer 54% conserved"
     repeat_region   32243. .32434
                     /note="6 copies 32 mer 64% conserved"
     repeat_region   32271. .32535
                     /note="5 copies 53 mer 64% conserved"
     repeat_region   32883. .32952
                     /note="2 copies 35 mer 92% conserved"
     repeat_region   33336. .33475
                     /note="4 copies 35 mer 70% conserved"
     repeat_region   33367. .33893

  Query Match             93.8%;  Score 93.8;  DB 9;  Length 132933;
  Best Local Similarity   97.9%;  Pred. No. 7.2e-18;
  Matches   95;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 CTTTGGCAGTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCC 63
           | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   65553 CCTTTGCAGTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCC 65612

Qy      64 ACGAGGAACAGAACTCTGTGGTGTGCTCCTGCGCCCG 100
           |||||||||||||||||||||||||||||||||||||
Db   65613 ACGAGGAACAGAACTCTGTGGTGTGCTCCTGCGCCCG 65649


RESULT 4
AB005892
LOCUS       AB005892                 350 bp    DNA     linear   PRI 20-JUN-2001
DEFINITION  Homo sapiens gene for coagulation factor X, partial cds.
ACCESSION   AB005892
VERSION     AB005892.1  GI:2281790
KEYWORDS    coagulation factor X; F10.
SOURCE      Homo sapiens mononuclear cell DNA.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Hayashi,T., Yahagi,A., Suzuki,K., Sato,S. and Sasaki,H.
  TITLE     Molecular abnormality observed in a patient with coagulation factor
            X (FX) deficiency: a novel three-base-pair (CTT) deletion within
            the polypyrimidine tract of the FX intron D
  JOURNAL   Br. J. Haematol. 102, 923-928 (1998)
REFERENCE   2  (bases 1 to 350)
  AUTHORS   Hayashi,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-JUL-1997) Tomohiro Hayashi, Yamagata University
            School of Medicine; 2-2-2 Iida-Nishi, Yamagata, Yamagata 990-23,
            Japan (E-mail:thayashi@med.id.yamagata-u.ac.jp,
            Tel:+81-236-28-5316, Fax:+81-236-28-5318)
FEATURES             Location/Qualifiers
     source          1. .350
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="13"
                     /map="13q32-13qter"
                     /cell_type="mononuclear cell"
     intron          1. .76
                     /note="intron D"
                     /number=4
     misc_difference 71^72
                     /note="Base delition has occurred between nucleotide 71
                     and 72"
                     /replace="ctt"
     gene            77. .208
                     /gene="F10"
     CDS             <77. .>208
                     /gene="F10"
                     /codon_start=1
                     /product="coagulation factor X"
                     /protein_id="BAA21634.1"
                     /db_xref="GI:2281791"
                     /translation="SHGSSAAWTTGTVTSSATRNRTLWCAPAPAGTPWLTTARPAFPQ
                     "
     exon            77. .208
                     /gene="F10"
                     /number=5
     intron          209. .350
                     /note="intron E"
                     /number=5
BASE COUNT       69 a    113 c    115 g     53 t
ORIGIN

  Query Match             92.0%;  Score 92;  DB 9;  Length 350;
  Best Local Similarity   100.0%;  Pred. No. 3.5e-17;
  Matches   92;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       9 GCAGTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAG 68
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      73 GCAGTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAG 132

Qy      69 GAACAGAACTCTGTGGTGTGCTCCTGCGCCCG 100
           ||||||||||||||||||||||||||||||||
Db     133 GAACAGAACTCTGTGGTGTGCTCCTGCGCCCG 164


RESULT 5
AR095306
LOCUS       AR095306                 1126 bp    DNA     linear   PAT 08-SEP-2000
DEFINITION   equence 27 from patent US 6004555.
ACCESSION   AR095306
VERSION     AR095306.1  GI:10023064
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1126)
  AUTHORS   Thorpe,P.E. and Edgington,T.S.
  TITLE     Methods for the specific coagulation of vasculature
  JOURNAL   Patent: US 6004555-A 27 21-DEC-1999;
FEATURES             Location/Qualifiers
     source          1. .1126
                     /organism="unknown"
BASE COUNT      269 a    341 c    342 g    174 t
ORIGIN

  Query Match             88.4%;  Score 88.4;  DB 6;  Length 1126;
  Best Local Similarity   98.9%;  Pred. No. 3.9e-16;
  Matches   89;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      11 AGTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 70
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      27 ATTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 86

Qy      71 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 100
           ||||||||||||||||||||||||||||||
Db      87 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 116


RESULT 6
AR103990
LOCUS       AR103990                 1126 bp    DNA     linear   PAT 14-FEB-2001
DEFINITION  Sequence 27 from patent US 6093399.
ACCESSION   AR103990
VERSION     AR103990.1  GI:12816698
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
```

```
REFERENCE   1  (bases 1 to 1126)
  AUTHORS   Thorpe,P.E. and Edgington,T.S.
  TITLE     Methods and compositions for the specific coagulation of
            vasculature
  JOURNAL   Patent: US 6093399-A 27 25-JUL-2000;
FEATURES             Location/Qualifiers
     source          1. .1126
                     /organism="unknown"
BASE COUNT      269 a    341 c    342 g    174 t
ORIGIN

  Query Match             88.4%;  Score 88.4;  DB 6;  Length 1126;
  Best Local Similarity   98.9%;  Pred. No. 3.9e-16;
  Matches   89;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      11 AGTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 70
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      27 ATTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 86

Qy      71 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 100
           ||||||||||||||||||||||||||||||
Db      87 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 116


RESULT 7
HUMFX
LOCUS       HUMFX                    1126 bp    mRNA    linear   PRI 08-NOV-1994
DEFINITION  Human factor X mRNA.
ACCESSION   K01886
VERSION     K01886.1  GI:182820
KEYWORDS    Stuart factor; factor X; serine protease.
SOURCE      Human liver, cDNA to mRNA, clone lambda-X-1137.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1126)
  AUTHORS   Leytus,S.P., Chung,D.W., Kisiel,W., Kurachi,K. and Davie,E.W.
  TITLE     Characterization of a cDNA coding for human factor X
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 81 (12), 3699-3702 (1984)
  MEDLINE   84222026
   PUBMED   6587384
COMMENT     In processing, factor X (Stuart factor) is converted to Xa by
            cleavage of a glycopeptide from the amino-terminal end of the heavy
            chain. It then acts as a serine protease in converting prothrombin
            to thrombin.
FEATURES             Location/Qualifiers
     source          1. .1126
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /map="13q34"
     gene            1. .1126
                     /gene="F10"
     mRNA            <1. .1126
                     /gene="F10"
                     /product="factor X mRNA"
     CDS             <1. .1116
                     /gene="F10"
                     /note="factor X precursor peptide"
                     /codon_start=1
                     /protein_id="AAA52486.1"
                     /db_xref="GI:182821"
                     /db_xref="GDB:G00-119-890"
                     /translation="GFEGKNCELFTRKLCSLDNGDCDQFCHEEQNSVVCSCARGYTLA
                     DNGKACIPTGPYPCGKQTLERRKRSVAQATSSSGEAPDSITWKPYDAADLDPTENPFD
                     LLDFNQTQPERGDNNLTRIVGGQECKDGECPWQALLINEENEGFCGGTILSEFYILTA
                     AHCLYQAKRFEGDRNTEQEEGGEAVHEVEVVIKHNRFTKETYDFDIAVLRLKTPITFR
                     MNVAPACLPERDWAESTLMTQKTGIVSGFGRTHEKGRQSTRLKMLEVPYVDRNSCKLS
                     SSFIITQNMFCAGYDTKQEDACQGDSGGPHVTRFKDTYFVTGIVSWGEGCARKGKYGI
                     YTKVTAFLKWIDRSMKTRGLPKAKSHAPEVITSSPLK"
     mat_peptide     <1. .195
                     /gene="F10"
                     /product="factor X light chain"
     mat_peptide     205. .1113
                     /gene="F10"
                     /product="factor X heavy chain"
     mat_peptide     361. .1113
                     /gene="F10"
                     /product="factor Xa heavy chain"
BASE COUNT      269 a    341 c    342 g    174 t
ORIGIN      5 bp upstream of TaqI site.

  Query Match             88.4%;  Score 88.4;  DB 9;  Length 1126;
  Best Local Similarity   98.9%;  Pred. No. 3.9e-16;
  Matches   89;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      11 AGTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 70
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      27 ATTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 86

Qy      71 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 100
           ||||||||||||||||||||||||||||||
Db      87 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 116


RESULT 8
A93124
LOCUS       A93124                   1404 bp    DNA     linear   PAT 22-JAN-2000
DEFINITION  Sequence 15 from Patent WO9747737.
ACCESSION   A93124
VERSION     A93124.1  GI:6741514
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 1404)
  AUTHORS   Kopetzki,E. and Hopfner,K.
  TITLE     RECOMBINANT BLOOD-COAGULATION PROTEASES
  JOURNAL   Patent: WO 9747737-A 15 18-DEC-1997;
            KOPETZKI ERHARD (DE); BOEHRINGER MANNHEIM GMBH (DE)
FEATURES             Location/Qualifiers
     source          1. .1404
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT      356 a    402 c    425 g    221 t
ORIGIN

  Query Match             88.4%;  Score 88.4;  DB 6;  Length 1404;
  Best Local Similarity   98.9%;  Pred. No. 3.8e-16;
  Matches   89;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0

Qy      11 AGTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 70
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     315 ATTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 374

Qy      71 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 100
           ||||||||||||||||||||||||||||||
Db     375 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 404


RESULT 9
HUMCFX
LOCUS       HUMCFX                   1414 bp    mRNA    linear   PRI 01-NOV-1994
DEFINITION  Human blood-coagulation factor X mRNA, complete cds.
ACCESSION   M22613
VERSION     M22613.1  GI:180335
KEYWORDS    coagulation factor X.
SOURCE      Human liver, cDNA to mRNA, clone pKT218.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1414)
  AUTHORS   Kaul,R.K., Hildebrand,B., Roberts,S. and Jagadeeswaran,P.
  TITLE     Isolation and characterization of human blood-coagulation factor X
            cDNA
  JOURNAL   Gene 41 (2-3), 311-314 (1986)
  MEDLINE   86221713
```

```
   PUBMED   3011603
FEATURES             Location/Qualifiers
     source          1. .1414
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /map="13q34"
     gene            1. .1414
                     /gene="F10"
     mRNA            <1. .1414
                     /gene="F10"
                     /product="coagulation factor X mRNA"
     CDS             <1. .1404
                     /gene="F10"
                     /note="coagulation factor X precursor"
                     /codon_start=1
                     /protein_id="AAA51984.1"
                     /db_xref="GI:180336"
                     /db_xref="GDB:G00-119-890"
                     /translation="LLGESLFIRREQANNILARVTRANSFLEEMKKGHLERECMEETC
                     SYEEAREVFEDSDKTNEFWNKYKDGDQCETSPCQNQGKCKDGLGEYTCTCLEGFEGKN
                     CELFTRKLCSLDNGDCDQFCHEEQNSVVCSCARGYTLADNGKACIPTGPYPCGKQTLE
                     RRKRSVAQATSSSGEAPDSITWKPYDAADLDPTENPFDLLDFNQTQPERGDNNLTRIV
                     GGQECKDGECPWQALLINEENEGFCGGTILSEFYILTAAHCLYQAKRFEGDRNTEQEE
                     GGEAVHEVEVVIKHNRFTKETYDFDIAVLRLKTPITFRMNVAPACLPERDWAESTLMT
                     QKTGIVSGFGRTHEKGRQSTRLKMLEVPYVDRNSCKLSSSFIITQNMFCAGYDTKQED
                     ACQGDSGGPHVTRFKDTYFVTGIVSWGEGCARKGKYGIYTKVTAFLKWIDRSMKTRGL
                     PKAKSHAPEVITSSPLK"
     sig_peptide     <1. .66
                     /gene="F10"
                     /note="coagulation factor X signal peptide"
     mat_peptide     67. .483
                     /gene="F10"
                     /product="coagulation factor X light chain"
     mat_peptide     493. .1401
                     /gene="F10"
                     /product="coagulation factor X heavy chain"
     mat_peptide     493. .648
                     /gene="F10"
                     /product="coagulation factor X activation peptide"
BASE COUNT      358 a    407 c    426 g    223 t
ORIGIN

  Query Match             88.4%;  Score 88.4;  DB 9;  Length 1414;
  Best Local Similarity   98.9%;  Pred. No. 3.8e-16;
  Matches   89;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      11 AGTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 70
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     315 ATTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 374

Qy      71 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 100
           ||||||||||||||||||||||||||||||
Db     375 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 404


RESULT 10
HUMFXM
LOCUS       HUMFXM                  1443 bp    mRNA    linear   PRI 08-NOV-1994
DEFINITION  Human factor X mRNA, partial signal pept and complete mature pept
            cds.
ACCESSION   K03194
VERSION     K03194.1  GI:182840
KEYWORDS    blood coagulation factor; factor X.
SOURCE      Human adult liver, cDNA to mRNA, clones pcHX[5,8,14].
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2)
  AUTHORS   MacGillivray,R.T.A.
  JOURNAL   Unpublished (1985)
REFERENCE   2  (bases 3 to 1443)
  AUTHORS   Fung,M.R., Hay,C.W. and MacGillivray,R.T.
  TITLE     Characterization of an almost full-length cDNA coding for human
            blood coagulation factor X
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 82 (11), 3591-3595 (1985)
  MEDLINE   85216545
   PUBMED   2582420
COMMENT     During conversion of factor X to factor X-a, a glycopeptide of 52
            amino acids (encoded by positions 513-668 in this sequence) is
            released.
            A polyadenylation signal is located at position 1424-1429. This
            sequence was kindly submitted over electronic mail by
            R.T.A.MacGillivray (23-SEP-1985).
FEATURES             Location/Qualifiers
     source          1. .1443
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /map="13q34"
     gene            1. .1443
                     /gene="F10"
     mRNA            <1. .1443
                     /gene="F10"
                     /product="X mRNA"
     CDS             1. .1433
                     /partial
                     /gene="F10"
                     /note="factor X prepeptide"
                     /codon_start=3
                     /protein_id="AAA52490.1"
                     /db_xref="GI:182841"
                     /db_xref="GDB:G00-119-890"
                     /translation="SLAGLLLLGESLFIRREQANNILARVTRANSFLEEMKKGHLERE
                     CMEETCSYEEAREVFEDSDKTNEFWNKYKDGDQCETSPCQNQGKCKDGLGEYTCTCLE
                     GFEGKNCELFTRKLCSLDNGDCDQFCHEEQNSVVCSCARGYTLADNGKACIPTGPYPC
                     GKQTLERRKRSVAQATSSSGEAPDSITWKPYDAADLDPTENPFDLLDFNQTQPERGDN
                     NLTRIVGGQECKDGECPWQALLINEENEGFCGGTILSEFYILTAAHCLYQAKRFKVRV
                     GDRNTEQEEGGEAVHEVEVVIKHNRFTKETYDFDIAVLRLKTPITFRMNVAPACLPER
                     DWAESTLMTQKTGIVSGFGRTHEKGRQSTRLKMLEVPYVDRNSCKLSSSFIITQNMFC
                     AGYDTKQEDACQGDSGGPHVTRFKDTYFVTGIVSWGESCARKGKYGIYTKVTAFLKWI
                     DRSMKTRGLPKAKSHAPEVITSSPLK"
     sig_peptide     <1. .86
                     /gen  "F10"
                     /note="factor X signal peptide"
     mat_peptide     87. .503
                     /gene="F10"
                     /product="factor X light chain"
     mat_peptide     513. .1430
                     /gene="F10"
                     /product="factor X heavy chain"
     variation       1117
                     /gene="F10"
                     /note="a in pcHX8; t in pcHX5"
BASE COUNT      361 a    416 c    435 g    231 t
ORIGIN      24 bp upstream of AvaI site.

  Query Match             88.4%;  Score 88.4;  DB 9;  Length 1443;
  Best Local Similarity   98.9%;  Pred. No. 3.8e-16;
  Matches   89;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      11 AGTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 70
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     335 ATTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 394

Qy      71 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 100
           ||||||||||||||||||||||||||||||
Db     395 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 424


RESULT 11
A86859
LOCUS       A86859                  1467 bp    DNA     linear   PAT 22-JAN-2000
DEFINITION  Sequence 43 from Patent WO9838318.
ACCESSION   A86859
VERSION     A86859.1  GI:6735650
KEYWORDS    .
SOURCE      unidentified.
```

```
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 1467)
  AUTHORS   Falkner,F. and Himmelspach,M.
  TITLE     FACTOR X DELETION MUTANTS AND ANALOGUES THEREOF
  JOURNAL   Patent: WO 9838318-A 43 03-SEP-1998;
            FALKNER FALKO GUENTER (AT); HIMMELSPACH MICHELE (AT)
FEATURES             Location/Qualifiers
     source          1. .1467
                     /organism="unidentified"
                     /db_xref="taxon:32644"
     CDS             1. .1467
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAB69367.1"
                     /db_xref="GI:6735651"
                     /translation="MGRPLHLVLLSASLAGLLLLGESLFIRREQANNILARVTRANSF
                     LEEMKKGHLERECMEETCSYEEAREVFEDSDKTNEFWNKYKDGDQCETSPCQNQGKCK
                     DGLGEYTCTCLEGFEGKNCELFTRKLCSLDNGDCDQFCHEEQNSVVCSCARGYTLADN
                     GKACIPTGPYPCGKQTLERRKRSVAQATSSSGEAPDSITWKPYDAADLDPTENPFDLL
                     DFNQTQPERGDNNLTRIVGGQECKDGECPWQALLINEENEGFCGGTILSEFYILTAAH
                     CLYQAKRFKVRVGDRNTEQEEGGEAVHEVEVVIKHNRFTKETYDFDIAVLRLKTPITF
                     RMNVAPACLPERDWAESTLMTQKTGIVSGFGRTHEKGRQSTRLKMLEVPYVDRNSCKL
                     SSSFIITQNMFCAGYDTKQEDACQGDSGGPHVTRFKDTYFVTGIVSWGESCARKGKYG
                     IYTKVTAFLKWIDRSMKTRGLPKAKSHAPEVITSSPLK"
BASE COUNT      363 a    424 c    444 g    236 t
ORIGIN

  Query Match             88.4%;  Score 88.4;  DB 6;  Length 1467;
  Best Local Similarity   98.9%;  Pred. No. 3.8e-16;
  Matches   89;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      11 AGTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 70
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     369 ATTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 428

Qy      71 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 100
           ||||||||||||||||||||||||||||||
Db     429 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 458


RESULT 12
A86886
LOCUS       A86886                 1467 bp    DNA     linear   PAT 22-JAN-2000
DEFINITION  Sequence 26 from Patent WO9838317.
ACCESSION   A86886
VERSION     A86886.1  GI:6735677
KEYWORDS    .
  URCE      unidentified.
  RGANISM   unidentified
            unclassified.
REFERENCE   1  (bases 1 to 1467)
  AUTHORS   Himmelspach,M. and Eibl,J.
  TITLE     FACTOR X ANALOGUES WITH A MODIFIED PROTEASE CLEAVAGE SITE
  JOURNAL   Patent: WO 9838317-A 26 03-SEP-1998;
            HIMMELSPACH MICHELE (AT); EIBL JOHANN (AT)
FEATURES             Location/Qualifiers
     source          1. .1467
                     /organism="unidentified"
                     /db_xref="taxon:32644"
     CDS             1. .1467
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAB69368.1"
                     /db_xref="GI:6735678"
                     /translation="MGRPLHLVLLSASLAGLLLLGESLFIRREQANNILARVTRANSF
                     LEEMKKGHLERECMEETCSYEEAREVFEDSDKTNEFWNKYKDGDQCETSPCQNQGKCK
                     DGLGEYTCTCLEGFEGKNCELFTRKLCSLDNGDCDQFCHEEQNSVVCSCARGYTLADN
                     GKACIPTGPYPCGKQTLERRKRSVAQATSSSGEAPDSITWKPYDAADLDPTENPFDLL
                     DFNQTQPERGDNNLTRIVGGQECKDGECPWQALLINEENEGFCGGTILSEFYILTAAH
                     CLYQAKRFKVRVGDRNTEQEEGGEAVHEVEVVIKHNRFTKETYDFDIAVLRLKTPITF
                     RMNVAPACLPERDWAESTLMTQKTGIVSGFGRTHEKGRQSTRLKMLEVPYVDRNSCKL
                     SSSFIITQNMFCAGYDTKQEDACQGDSGGPHVTRFKDTYFVTGIVSWGESCARKGKYG
                     IYTKVTAFLKWIDRSMKTRGLPKAKSHAPEVITSSPLK"
BASE COUNT      363 a    424 c    444 g    236 t
ORIGIN

  Query Match             88.4%;  Score 88.4;  DB 6;  Length 1467;
  Best Local Similarity   98.9%;  Pred. No. 3.8e-16;
  Matches   89;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      11 AGTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 70
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     369 ATTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 428

Qy      71 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 100
           ||||||||||||||||||||||||||||||
Db     429 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 458


RESULT 13
AX082959
LOCUS       AX082959               1467 bp    DNA     linear   PAT 28-FEB-2001
DEFINITION  Sequence 1 from Patent WO0110896.
ACCESSION   AX082959
VERSION     AX082959.1  GI:13184880
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1467)
  AUTHORS   Himmelspach,M. and Schlokat,U.
  TITLE     Factor x analog with an improved ability to be activated
  JOURNAL   Patent: WO 0110896-A 1 15-FEB-2001;
            Baxter Aktiengesellschaft (AT)
FEATURES             Location/Qualifiers
     source          1. .1467
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      363 a    424 c    444 g    236 t
ORIGIN

  Query Match             8.4%;  Score 88.4;  DB 6;  Length 1467;
  Best Local Similarity   98.9%;  Pred. No. 3.8e-16;
  Matches   89;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      11 AGTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 70
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     369 ATTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 428

Qy      71 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 100
           ||||||||||||||||||||||||||||||
Db     429 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 458


RESULT 14
AR024194
LOCUS       AR024194               1500 bp    DNA     linear   PAT 05-DEC-1998
DEFINITION  Sequence 4 from patent US 5795863.
ACCESSION   AR024194
VERSION     AR024194.1  GI:3977488
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1500)
  AUTHORS   Wolf,D.
  TITLE     Recombinant agents affecting thrombosis
  JOURNAL   Patent: US 5795863-A 4 18-AUG-1998;
FEATURES             Location/Qualifiers
     source          1. .1500
                     /organism="unknown"
BASE COUNT      390 a    425 c    441 g    244 t
ORIGIN
```

```
  Query Match             88.4%;  Score 88.4;  DB 6;  Length 1500;
  Best Local Similarity   98.9%;  Pred. No. 3.8e-16;
  Matches   89;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      11 AGTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 70
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     401 ATTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 460

Qy      71 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 100
           ||||||||||||||||||||||||||||||
Db     461 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 490


RESULT 15
HUMFACX
LOCUS       HUMFACX                 1507 bp    mRNA    linear   PRI 08-NOV-1994
DEFINITION  Human coagulation factor X (F10) mRNA, complete cds.
ACCESSION   M57285
VERSION     M57285.1  GI:182389
KEYWORDS    coagulation factor X.
SOURCE      Human, cDNA to mRNA.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1507)
  AUTHORS   Messier,T.L., Pittman,D.D., Long,G.L., Kaufman,R.J. and Church,W.R.
  TITLE     Cloning and expression in COS-1 cells of a full-length cDNA
            encoding human coagulation factor X
  JOURNAL   Gene 99 (2), 291-294 (1991)
  MEDLINE   91216473
   PUBMED   1902434
FEATURES             Location/Qualifiers
     source          1. .1507
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /map="13q34"
                     /tissue_type="liver"
     gene            1. .1507
                     /gene="F10"
     CDS             1. .1467
                     /gene="F10"
                     /EC_number="3.4.21.6"
                     /codon_start=1
                     /product="coagulation factor X"
                     /protein_id="AAA52421.1"
                     /db_xref="GI:182390"
                     /db_xref="GDB:G00-119-890"
                     /translation="MGRPLHLVLLSASLAGLLLLGESLFIRREQANNILARVTRANSF
                     LEEMKKGHLERECMEETCSYEEAREVFEDSDKTNEFWNKYKDGDQCETSPCQNQGKCK
                     DGLGEYTCTCLEGFEGKNCELFTRKLCSLDNGDCDQFCHEEQNSVVCSCARGYTLADN
                     GKACIPTGPYPCGKQTLERRKRSVAQATSSSGEAPDSITWKPYDAADLDPTENPFDLL
                     DFNQTQPERGDNNLTRIVGGQECKDGECPWQALLINEENEGFCGGTILSEFYILTAAH
                     CLYQAKRFKVRVGDRNTEQEEGGEAVHEVEVVIKHNRFTKETYDFDIAVLRLKTPITF
                     RMNVAPACLPERDWAESTLMTQKTGIVSGFGRTHEKGRQSTRLKMLEVPYVDRNSCKL
                     SSSFIITQNMFCAGYDTKQEDACQGDSGGPHVTRFKDTYFVTGIVSWGEGCARKGKYG
                     IYTKVTAFLKWIDRSMKTRGLPKAKSHAPEVITSSPLK"
     misc_feature    1004. .1060
                     /gene="F10"
                     /note="putative VECTOR sequence Bacteriophage lambda
                     (J02459); putative"
BASE COUNT      394 a    429 c    446 g    238 t
ORIGIN

  Query Match             88.4%;  Score 88.4;  DB 9;  Length 1507;
  Best Local Similarity   98.9%;  Pred. No. 3.8e-16;
  Matches   89;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      11 AGTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 70
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     369 ATTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 428

Qy      71 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 100
           ||||||||||||||||||||||||||||||
Db     429 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 458


Search completed: January 15, 2003, 20:35:16
Job time : 1355 secs
```

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: January 15, 2003, 17:58:10 ; Search time 31 Seconds
(without alignments)
989.280 Million cell updates/sec

Title: L00394_COPY_1_100
Perfect score: 100
Sequence: 1 CTCCTTTGGCAGTCACACGG.....GTGGTGTGCTCCTGCGCCCG 100

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	88.4	88.4	1126	2	US-08-479-733A-27	Sequence 27, Appl
	2	88.4	88.4	1126	3	US-08-487-427-27	Sequence 27, Appl
	3	88.4	88.4	1126	3	US-08-479-727A-27	Sequence 27, Appl
	4	88.4	88.4	1126	3	US-08-482-369A-27	Sequence 27, Appl
	5	88.4	88.4	1126	5	PCT-US95-07439-27	Sequence 27, Appl
	6	88.4	88.4	1404	4	US-09-202-101-15	Sequence 15, Appl
	7	88.4	88.4	1500	1	US-08-487-037-4	Sequence 4, Appli
	8	52.4	52.4	1554	1	US-08-469-486-1	Sequence 1, Appli
	9	52.4	52.4	1554	2	US-08-469-658-1	Sequence 1, Appli
	10	30.8	30.8	2461	1	US-08-282-141-1	Sequence 1, Appli
	11	28.2	28.2	22306	4	US-09-453-702B-251	Sequence 251, App
c	12	28.2	28.2	46819	4	US-09-453-702B-72	Sequence 72, Appl
c	13	26.8	26.8	4403765	4	US-09-103-840A-2	Sequence 2, Appli
c	14	26.8	26.8	4411529	4	US-09-103-840A-1	Sequence 1, Appli
c	15	26.4	26.4	3156	4	US-09-284-819-8	Sequence 8, Appli
	16	25.8	25.8	10095	3	US-08-822-586-45	Sequence 45, Appl
	17	25.4	25.4	286	2	US-08-675-508-21	Sequence 21, Appl
	18	25.4	25.4	1409	4	US-09-338-907-72	Sequence 72, Appl
	19	25.4	25.4	1409	4	US-09-338-907-184	Sequence 184, App
	20	25.4	25.4	1409	4	US-09-218-207-72	Sequence 72, Appl
	21	25.4	25.4	1409	4	US-09-218-207-184	Sequence 184, App
	22	25.4	25.4	4550	4	US-09-338-907-182	Sequence 182, App
	23	25.4	25.4	4550	4	US-09-218-207-182	Sequence 182, App
	24	25.2	25.2	1866	4	US-09-615-192A-103	Sequence 103, App
	25	25.2	25.2	2013	4	US-09-615-192A-404	Sequence 404, App
c	26	25.2	25.2	6677	4	US-08-939-366-27	Sequence 27, Appl
c	27	25.2	25.2	6677	4	US-09-467-997-6	Sequence 6, Appli
	28	25.2	25.2	10079	2	US-08-476-866-20	Sequence 20, Appl
	29	25	25.0	1386	2	US-08-756-506-3	Sequence 3, Appli
	30	25	25.0	1387	6	5270178-1	Patent No. 5270178
	31	25	25.0	1755	6	5225537-1	Patent No. 5225537
c	32	25	25.0	2373	4	US-08-980-080-1	Sequence 1, Appli
c	33	25	25.0	3003	1	US-08-434-730-15	Sequence 15, Appl
c	34	25	25.0	4495	4	US-09-417-822-6	Sequence 6, Appli
c	35	25	25.0	4534	4	US-09-417-822-7	Sequence 7, Appli
c	36	25	25.0	4535	4	US-09-417-822-8	Sequence 8, Appli
c	37	25	25.0	8838	4	US-09-417-822-1	Sequence 1, Appli
	38	25	25.0	11725	2	US-08-756-506-1	Sequence 1, Appli
	39	24.8	24.8	288	2	US-08-675-508-23	Sequence 23, Appl
	40	24.8	24.8	494	2	US-08-675-508-4	Sequence 4, Appli
	41	24.8	24.8	998	4	US-09-203-939-1	Sequence 1, Appli
	42	24.8	24.8	998	4	US-09-251-835-1	Sequence 1, Appli
	43	24.8	24.8	998	4	US-09-318-503-1	Sequence 1, Appli
	44	24.8	24.8	998	4	US-09-038-261A-1	Sequence 1, Appli
c	45	24.8	24.8	1485	4	US-09-088-435-2	Sequence 2, Appli

ALIGNMENTS

```
RESULT 1
US-08-479-733A-27
; Sequence 27, Application US/08479733A
; Patent No. 5877289
;  GENERAL INFORMATION:
;    APPLICANT:  Thorpe, Philip E.
;    APPLICANT:  Edgington, Thomas S.
;    TITLE OF INVENTION:  Methods and Compositions for the
;    TITLE OF INVENTION:  Specific Coagulation of Vasculature
;    NUMBER OF SEQUENCES:  32
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  US
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/479,733A
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/273,567
;      FILING DATE:  11-JUL-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Parker, David L.
;      REGISTRATION NUMBER:  32,165
;      REFERENCE/DOCKET NUMBER:  UTSD:459/PAR
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  512/418-3000
;      TELEFAX:  512/474-7577
;      TELEX:  N/A
;  INFORMATION FOR SEQ ID NO:  27:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1126 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-479-733A-27

  Query Match             88.4%;  Score 88.4;  DB 2;  Length 1126;
  Best Local Similarity   98.9%;  Pred. No. 1e-19;
  Matches   89;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      11 AGTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 70
```

```
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      27 ATTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 86

Qy      71 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 100
           ||||||||||||||||||||||||||||||
Db      87 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 116


RESULT 2
US-08-487-427-27
; Sequence 27, Application US/08487427
; Patent No. 6004555
;  GENERAL INFORMATION:
;    APPLICANT:  Thorpe, Philip E.
;    APPLICANT:  Edgington, Thomas S.
;    TITLE OF INVENTION:  Methods and Compositions for the
;    TITLE OF INVENTION:  Specific Coagulation of Vasculature
;    NUMBER OF SEQUENCES:  27
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  US
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/487,427
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/273,567
;      FILING DATE:  11-JUN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Parker, David L.
;      REGISTRATION NUMBER:  32,165
;      REFERENCE/DOCKET NUMBER:  UTSD:457/PAR
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  512/418-3000
;      TELEFAX:  512/474-7577
;      TELEX:  N/A
;  INFORMATION FOR SEQ ID NO:  27:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1126 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-487-427-27

  Query Match             88.4%;  Score 88.4;  DB 3;  Length 1126;
  Best Local Similarity   98.9%;  Pred. No. 1e-19;
  Matches   89;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      11 AGTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 70
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      27 ATTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 86

Qy      71 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 100
           ||||||||||||||||||||||||||||||
Db      87 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 116


RESULT 3
US-08-479-727A-27
; Sequence 27, Application US/08479727A
; Patent No. 6036955
;  GENERAL INFORMATION:
;    APPLICANT:  Thorpe, Philip E.
;    TITLE OF INVENTION:  Methods and Compositions for the
;    TITLE OF INVENTION:  Specific Coagulation of Vasculature
;    NUMBER OF SEQUENCES:  32
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  US
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/479,727A
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/273,567
;      FILING DATE:  11-JUL-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Parker, David L.
;      REGISTRATION NUMBER:  32,165
;      REFERENCE/DOCKET NUMBER:  UTSD:461/PAR
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  512/418-3000
;      TELEFAX:  512/474-7577
;      TELEX:  N/A
;  INFORMATION FOR SEQ ID NO:  27:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1126 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-479-727A-27

  Query Match             88.4%;  Score 88.4;  DB 3;  Length 1126;
  Best Local Similarity   98.9%;  Pred. No. 1e-19;
  Matches   89;  Conservative    0;  Mismatches    ;  Indels    0;  Gaps    0;

Qy      11 AGTCACAC GGAAGCTCTGCAGCCTGGACAACGGGGACTGTG CCAGTTCTGCCACGAGGA 70
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      27 ATTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 86

Qy      71 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 100
           ||||||||||||||||||||||||||||||
Db      87 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 116


RESULT 4
US-08-482-369A-27
; Sequence 27, Application US/08482369A
; Patent No. 6093399
;  GENERAL INFORMATION:
;    APPLICANT:  Thorpe, Philip E.
;    APPLICANT:  Edgington, Thomas S.
;    TITLE OF INVENTION:  Methods and Compositions for the Specific
;    TITLE OF INVENTION:  Coagulation of Vasculature
;    NUMBER OF SEQUENCES:  32
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  US
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
```

```
;       APPLICATION NUMBER:  US/08/482,369A
;       FILING DATE:  07-JUN-1995
;       CLASSIFICATION:  424
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/273,567
;       FILING DATE:  11-JUL-1994
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Parker, David L.
;       REGISTRATION NUMBER:  32,165
;       REFERENCE/DOCKET NUMBER:  UTSD:433/PAR
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  512/418-3000
;       TELEFAX:  512/474-7577
;       TELEX:  N/A
;   INFORMATION FOR SEQ ID NO:  27:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1126 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;-08-482-369A-27

  Query Match             88.4%;  Score 88.4;  DB 3;  Length 1126;
  Best Local Similarity   98.9%;  Pred. No. 1e-19;
  Matches   89;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      11 AGTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 70
           | |||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db      27 ATTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACT TGACCAGTTCTGCCACGAGGA 86

Qy      71 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 100
           ||||||||||||||||||||||||||||||
Db      87 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 116


RESULT 5
PCT-US95-07439-27
; Sequence 27, Application PC/TUS9507439
;  GENERAL INFORMATION:
;    APPLICANT:
;    APPLICANT:  NAME:     BOARD OF REGENTS, THE UNIVERSITY OF
;    APPLICANT:  TEXAS SYSTEM
;    APPLICANT:  STREET:   201 West 7th Street
;    APPLICANT:  CITY:     Austin
;    APPLICANT:  STATE:    Texas
;    APPLICANT:  COUNTRY:  United States of America
;    APPLICANT:  POSTAL CODE:  78701
;    APPLICANT:  TELEPHONE NO: (512)499-4462
;    APPLICANT:  TELEFAX:  (512)499-4523
;    APPLICANT:  NAME:     THE SCRIPPS RESEARCH INSTITUTE
;    APPLICANT:  STREET:   10666 North Torrey Pines Road
;    APPLICANT:  CITY:     LaJolla
;    APPLICANT:  STATE:    California
;    APPLICANT:  COUNTRY:  United States of America
;    APPLICANT:  POSTAL CODE:  92037
;    TITLE OF INVENTION:  METHODS AND COMPOSITIONS
;    TITLE OF INVENTION:  FOR THE SPECIFIC
;    TITLE OF INVENTION:  COAGULATION OF VASCULATURE
;    NUMBER OF SEQUENCES:  27
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P. O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  USA
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS, ASCII
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/07439
;      FILING DATE:  Concurrently herewith
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/273,567
;      FILING DATE:  11-JUN 1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  PARKER, DAVID L.
;      REGISTRATION NUMBER:  32,165
;      REFERENCE/DOCKET NUMBER:  UTFD433P--
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (512) 418-3000
;      TELEFAX:  (713) 789-2679
;      TELEX:  79-0924
;  INFORMATION FOR SEQ ID NO:  27:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1126 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
PCT-US95-07439-27

  Query Match             88.4%;  Score 88.4;  DB 5;  Length 1126;
  Best Local Similarity   98.9%;  Pred. No. 1e-19;
  Matches   89;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      11 AGTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 70
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      27 ATTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 86

Qy      71 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 100
           ||||||||||||||||||||||||||||||
Db      87 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 116


RESULT 6
US-09-202-101-15
; Sequence 15, Application US/09202101
; Patent No. 6277618
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  Recombinant blood-coagulation proteases
;    NUMBER OF SEQUENCES:  17
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30B (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/202,101
;      FILING DATE:
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1404 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-202-101-15

  Query Match             88.4%;  Score 88.4;  DB 4;  Length 1404;
  Best Local Similarity   98.9%;  Pred. No. 1.1e-19;
  Matches   89;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      11 AGTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 70
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     315 ATTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 374

Qy      71 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 100
           ||||||||||||||||||||||||||||||
Db     375 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 404


RESULT 7
US-08-487-037-4
```

```
; Sequence 4, Application US/08487037
; Patent No. 5795863
;  GENERAL INFORMATION:
;    APPLICANT:  Wolf, David L.
;    TITLE OF INVENTION:  RECOMBINANT AGENTS AFFECTING THROMBOSIS
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  2000 Pennsylvania Avenue, NW
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20006-1812
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/487,037
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Adler, Reid G.
;      REGISTRATION NUMBER:  30,988
;      REFERENCE/DOCKET NUMBER:  2803-0002.02
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 887-1500
;      TELEFAX:  (202) 887-0763
;      TELEX:  90-4030
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1500 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-487-037-4

  Query Match             88.4%;  Score 88.4;  DB 1;  Length 1500;
  Best Local Similarity   98.9%;  Pred. No. 1.1e-19;
  Matches   89;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      11 AGTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 70
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     401 ATTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGA 460

Qy      71 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 100
           |||||||||||||||||||||||||||||
Db     461 ACAGAACTCTGTGGTGTGCTCCTGCGCCCG 490


RESULT 8
US-08-469-486-1
; Sequence 1, Application US/08469486
; Patent No. 5739281
;  GENERAL INFORMATION:
;    APPLICANT:  Thoegersen, Hans Christian
;    APPLICANT:  Holtet, Thor Las
;    APPLICANT:  Etzerodt, Michael
;    TITLE OF INVENTION:  Improved method for the refolding of
;    TITLE OF INVENTION:  proteins
;    NUMBER OF SEQUENCES:  58
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version
;      SOFTWARE:  #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/469,486
;      FILING DATE:
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/192,060
;      FILING DATE:  February 4, 1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Paul T. Clark
;      REGISTRATION NUMBER:  30,162
;      REFERENCE/DOCKET NUMBER:  06363/002001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617 542 5070
;      TELEFAX:  617 542 8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1554 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  YES
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Bos taurus
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  76..1551
US-08-469-486-1

  Query Match             52.4%;  Score 52.4;  DB 1;  Length 1554;
  Best Local Similarity   75.6%;  Pred. No. 4.4e-08;
  Matches   65;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy      14 CACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGAACA 73
           ||| ||  || |||||||||||||||| || | ||| ||||||||||||   ||||| |
Db     447 CACGCGTGAGATCTGCAGCCTGGACAATGGAGGCTGCGACCAGTTCTGCAGGGAGGAGCG 506

Qy      74 GAACTCTGTGGTGTGCTCCTGCGCCC 99
           | |   |||  |||||||||||| |
Db     507 CAGCGAGGTGCGGTGCTCCTGCGCGC 532


RESULT 9
US-08-469-658-1
; Sequence 1, Application US/08469658
; Patent No. 5917018
;  GENERAL INFORMATION:
;    APPLICANT:  Th egersen, Hans Christian
;    APPLICANT:  Holtet, Thor Las
;    APPLICANT:  Etzerodt, Michael
;    TITLE OF INVENTION:  IMPROVED METHOD FOR THE REFOLDING OF
;    TITLE OF INVENTION:  PROTEINS
;    NUMBER OF SEQUENCES:  58
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version
;      SOFTWARE:  #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/469,658
;      FILING DATE:  June 5, 1995
```

```
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/192,060
;      FILING DATE:  February 4, 1994
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Paul T. Clark
;      REGISTRATION NUMBER:  30,162
;      REFERENCE/DOCKET NUMBER:  06363/002002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617 542 5070
;      TELEFAX:  617 542 8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1554 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
     MOLECULE TYPE:  cDNA
     HYPOTHETICAL:  YES
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Bos taurus
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  76..1551
US-08-469-658-1

  Query Match             52.4%;  Score 52.4;  DB 2;  Length 1554;
  Best Local Similarity   75.6%;  Pred. No. 4.4e-08;
  Matches   65;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy      14 CACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGAACA 73
           ||| ||  || ||||||||||||||||| || | ||| ||||||||||||   ||||| |
Db     447 CACGCGTGAGATCTGCAGCCTGGACAATGGAGGCTGCGACCAGTTCTGCAGGGAGGAGCG 506

Qy      74 GAACTCTGTGGTGTGCTCCTGCGCCC 99
            | |   |||  |||||||||||| |
Db     507 CAGCGAGGTGCGGTGCTCCTGCGCGC 532


RESULT 10
US-08-282-141-1
; Sequence 1, Application US/08282141
; Patent No. 5538861
;  GENERAL INFORMATION:
;    APPLICANT:  Schneider, Claudio
;    APPLICANT:  Varnum, Brian
     APPLICANT:  Avanzi, Giancarlo
;    APPLICANT:  Brancolini, Claudio
;    APPLICANT:  Manfioletti, Guidalberto
;    TITLE OF INVENTION:  Stimulating Factor for the AXL Receptor
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Amgen Inc.
;      STREET:  1840 Dehavilland Drive
;      CITY:  Thousand Oaks
;      STATE:  California
;      COUNTRY:  United States
;      ZIP:  91320-1789
;    COMPUTER REA ABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/282,141
;      FILING DATE:
;      CLASSIFICATION:  435
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2461 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  cDNA
US-08-282-141-1

  Query Match             30.8%;  Score 30.8;  DB 1;  Length 2461;
  Best Local Similarity   63.5%;  Pred. No. 0.44;
  Matches   47;  Conservative    0;  Mismatches   27;  Indels    0;  Gaps    0;

Qy      27 TGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGAACAGAACTCTGTGGTG 86
           ||||||| ||| ||||||| |||   |||| |||||||| |  | | |       |
Db     612 TGCAGCCAGGAGAACGGGGGCTGCCTCCAGATCTGCCACAACAAGCCGGGTAGCTTCCAC 671

Qy      87 TGCTCCTGCGCCCG 100
           || ||||||  | |
Db     672 TGTTCCTGCCACAG 685


RESULT 11
US-09-453-702B-251
; Sequence 251, Application US/09453702B
; Patent No. 6365723
;  GENERAL INFORMATION:
;       APPLICANT: Blattner, Frederick R.
;                  Burland,    Valerie
;                  Perna,  Nicole T.
;                  Plunkett, Guy
;                  Welch,  Rod
;       TITLE OF INVENTION: No. 6365723el Sequences of E. coli O157
;       NUMBER OF SEQUENCES: 265
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Quarles & Brady
;            STREET: 1 South Pinckney Street
;            CITY: Madison
;            STATE: WI
;            COUNTRY: US
;            ZIP: 53701-2113
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Diskette, 3.50 inch. 1.44Mb storage
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: Word Perfect 8.0
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/453,702B
;            FILING DATE: 03-Dec-1999
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 60/110,955
;            FILING DATE: 04-DEC-1998
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Seay, Nicholas J.
;            REGISTRATION NUMBER: 27386
;            REFERENCE/DOCKET NUMBER: 960296.95017
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (608) 251-5000
;            TELEFAX: (608) 251-9166
;  INFORMATION FOR SEQ ID NO: 251:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 22306
;            TYPE: nucleic acid
;            STRANDEDNESS: double
;            TOPOLOGY: linear
;       MOLECULE TYPE: DNA  (genomic)
;       SEQUENCE DESCRIPTION: SEQ ID NO: 251:
US-09-453-702B-251

  Query Match             28.2%;  Score 28.2;  DB 4;  Length 22306;
  Best Local Similarity   61.6%;  Pred. No. 5.2;
  Matches   45;  Conservative    0;  Mismatches   28;  Indels    0;  Gaps    0;

Qy      18 CGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGAACAGAAC 77
           | || || || | | |||  || | || |||||| || |  | | | ||||| || |
```

```
Db  16880 CTGATGCGCTTCGGGCTGACCATCCGGAACTGTGTCCGGAACAGTCGCGAGGGACGGGTA 16939

Qy      78 TCTGTGGTGTGCT 90
           ||   ||||  ||
Db  16940 TCCCAGGTGGCCT 16952


RESULT 12
US-09-453-702B-72/c
; Sequence 72, Application US/09453702B
; Patent No. 6365723
;   GENERAL INFORMATION:
;        APPLICANT: Blattner, Frederick R.
;                   Burland,     Valerie
;                   Perna,  Nicole T.
;                   Plunkett, Guy
;                   Welch,  Rod
;        TITLE OF INVENTION: No. 6365723el Sequences of E. coli O157
;        NUMBER OF SEQUENCES: 265
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Quarles & Brady
;             STREET: 1 South Pinckney Street
;             CITY: Madison
;             STATE: WI
;             COUNTRY: US
;             ZIP: 53701-2113
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette, 3.50 inch. 1.44Mb storage
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: Word Perfect 8.0
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/453,702B
;             FILING DATE: 03-Dec-1999
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 60/110,955
;             FILING DATE: 04-DEC-1998
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Seay, Nicholas J.
;             REGISTRATION NUMBER: 27386
;             REFERENCE/DOCKET NUMBER: 960296.95017
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (608) 251-5000
;             TELEFAX: (608) 251-9166
;   INFORMATION FOR SEQ ID NO: 72:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 46819
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        MOLECULE TYPE: DNA  (genomic)
;        SEQUENCE DESCRIPTION: SEQ ID NO: 72:
US-09-453-702B-72

  Query Match             28.2%;  Score 28.2;  DB 4;  Length 46819;
  Best Local Similarity   61.6%;  Pred. No. 6.3;
  Matches   45;  Conservative    0;  Mismatches   28;  Indels    0;  Gaps    0;

Qy      18 CGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGAACAGAAC 77
           | || || || | | |||  || | || |||||| || |  | | | ||||| || |
Db  15108 CTGATGCGCTTCGGGCTGACCATCCGGAACTGTGTCCGGAACAGTCGCGAGGGACGGGTA 15049

Qy      78 TCTGTGGTGTGCT 90
           ||   ||||  ||
Db  15048 TCCCAGGTGGCCT 15036


RESULT 13
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
;  APPLICANT: FLEISCHMAN, Robert D.
;  APPLICANT:  WHITE, Owen R.
;  APPLICANT:  FRASER, Claire M.
;  APPLICANT:  VENTER, John C.
;  TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
;  TITLE OF INVENTION:  TUBERCULOSIS
;  FILE REFERENCE: 24366-20007.00
;  CURRENT APPLICATION NUMBER: US/09/103,840A
;  CURRENT FILING DATE:  1998-06-24
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 4403765
;   TYPE: DNA
;   ORGANISM: Mycobacterium tuberculosis
;   FEATURE:
;   OTHER INFORMATION: CDC 1551
;   OTHER INFORMATION: "n" bases at various positions throughout the sequence
;   OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

  Query Match             26.8%;  Score 26.8;  DB 4;  Length 4403765;
  Best Local Similarity   68.5%;  Pred. No. 38;
  Matches   37;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       6 TTGGCAGTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTC 59
           |||| | || ||| | |  ||| |  ||||||| ||| ||||| || |  | ||
Db 3915983 TTGGGAATCGCACCGGATATCTTCGACCTGGACGACGAGGACTATGCCGTGGTC 3915930


RESULT 14
US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
;  APPLICANT: FLEISCHMAN, Robert D.
;  APPLICANT:  WHITE, Owen R.
;  APPLICANT:  FRASER, Claire M.
;  APPLICANT:  VENTER, John C.
;  TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
;  TITLE OF INVENTION:  TUBERCULOSIS
;  FILE REFERENCE: 24366-20007.00
;  CURRENT APPLICATION NUMBER: US/09/103,840A
;  CURRENT FILING DATE:  1998-06-24
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 4411529
;   TYPE: DNA
;   ORGANISM: Mycobacterium tuberculosis
;   OTHER INFORMATION: H37Rv
US-09-103-840A-1

  Query Match             26.8%;  Score 26.8;  DB 4;  Length 4411529;
  Best Local Similarity   68.5%;  Pred. No. 38;
  Matches   37;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       6 TTGGCAGTCACACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTC 59
           |||| | || ||| | |  ||| |  ||||||| ||| ||||| || |  | ||
Db 3922204 TTGGGAATCGCACCGGATATCTTCGACCTGGACGACGAGGACTATGCCGTGGTC 3922151


RESULT 15
US-09-284-819-8/c
; Sequence 8, Application US/09284819
; Patent No. 6365712
; GENERAL INFORMATION:
;  APPLICANT: Kelly, Kathleen
;  APPLICANT: The Government of the United States of America
;  APPLICANT: as represented by The Secretary of the
;  APPLICANT: Department of Health and Human Services
;  TITLE OF INVENTION: Methods and Compositions for Inhibiting Inflammation
;  TITLE OF INVENTION: and Angiogenesis Comprising a Mammalian CD97 Alpha
```

```
;  TITLE OF INVENTION: Subunit
;  FILE REFERENCE: 015280-263100US
;  CURRENT APPLICATION NUMBER: US/09/284,819
;  CURRENT FILING DATE: 1999-08-20
;  EARLIER APPLICATION NUMBER: US 60/027,871
;  EARLIER FILING DATE: 1996-10-25
;  EARLIER APPLICATION NUMBER: WO PCT/US97/19772
;  EARLIER FILING DATE: 1997-10-24
;  NUMBER OF SEQ ID NOS: 21
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
;   LENGTH: 3156
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: full length pAT276 encoding CD97
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (49)..(2556)
    OTHER INFORMATION: CD97
3-09-284-819-8

  Query Match             26.4%;  Score 26.4;  DB 4;  Length 3156;
  Best Local Similarity   57.1%;  Pred. No. 12;
  Matches   48;  Conservative    0;  Mismatches   36;  Indels    0;  Gaps    0;

Qy      17 ACGGAAGCTCTGCAGCCTGGACAACGGGGACTGTGACCAGTTCTGCCACGAGGAACAGAA 76
           ||  ||  ||| || |||  ||||| | ||      || |  | |||| |  |  |||
Db    3093 ACACAACGTCTACAACCTTCACAACAGTGAGAACCTCCTGCCCGGCCAGGCAGGGCAGGC 3034

Qy      77 CTCTGTGGTGTGCTCCTGCGCCCG 100
           ||||||   |||  |||  | | |
Db    3033 CTCTGTACCGTGAGCCTCAGGCAG 3010


Search completed: January 15, 2003, 20:51:13
Job time : 933 secs
```

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: January 15, 2003, 20:35:40 ; Search time 32.5 Seconds
(without alignments)
1371.902 Million cell updates/sec

Title: L00396_COPY_1_100
Perfect score: 100
Sequence: 1 CGTCTGTCACAGGCCCTGCT.....AACGGCAGCCCACTGTCTCT 100

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 393868 seqs, 222934149 residues

Total number of hits satisfying chosen parameters: 787736

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	36	36.0	1047	9	US-09-808-602-100	Sequence 100, App
	2	34.2	34.2	1795	10	US-09-764-898-72	Sequence 72, Appl
	3	34.2	34.2	2063	9	US-10-063-547-111	Sequence 111, App
	4	34.2	34.2	2063	9	US-10-174-590-329	Sequence 329, App
	5	34.2	34.2	2063	9	US-10-176-758-329	Sequence 329, App
	6	34.2	34.2	2063	12	US-10-006-867-111	Sequence 111, App
	7	34.2	34.2	2063	12	US-10-052-586-329	Sequence 329, App
	8	34.2	34.2	2079	10	US-09-851-588-5	Sequence 5, Appli
	9	34.2	34.2	2081	10	US-09-851-588-7	Sequence 7, Appli
	10	30.2	30.2	396	10	US-09-960-352-10052	Sequence 10052, A
	11	30.2	30.2	420	10	US-09-960-352-6361	Sequence 6361, Ap
c	12	29.8	29.8	1674	9	US-09-738-626-1340	Sequence 1340, Ap
	13	29.6	29.6	1059	10	US-09-888-615-36	Sequence 36, Appl
	14	28.4	28.4	867	10	US-09-755-016-5	Sequence 5, Appli
	15	28.4	28.4	1047	10	US-09-755-016-3	Sequence 3, Appli
	16	28.4	28.4	1080	10	US-09-888-615-55	Sequence 55, Appl
	17	28.4	28.4	1286	10	US-09-755-016-7	Sequence 7, Appli
	18	28	28.0	424	10	US-09-960-352-5863	Sequence 5863, Ap
c	19	28	28.0	1383	10	US-09-841-880-3	Sequence 3, Appli
	20	28	28.0	1413	10	US-09-884-901-2	Sequence 2, Appli
	21	28	28.0	1548	10	US-09-150-811-6	Sequence 6, Appli
c	22	28	28.0	1937	10	US-09-841-880-1	Sequence 1, Appli
	23	28	28.0	2792	10	US-09-118-748-1	Sequence 1, Appli
c	24	28	28.0	32187	10	US-09-764-847-1550	Sequence 1550, Ap
c	25	28	28.0	32193	10	US-09-764-847-1549	Sequence 1549, Ap
	26	28	28.0	38059	10	US-09-880-107-2125	Sequence 2125, Ap
	27	27.8	27.8	3286	9	US-10-013-310-4	Sequence 4, Appli
	28	26.8	26.8	857	10	US-09-822-849A-158	Sequence 158, App
	29	26.6	26.6	224	10	US-09-728-446-1164	Sequence 1164, Ap
	30	26.6	26.6	987	10	US-09-888-615-32	Sequence 32, Appl
	31	26.6	26.6	1543	10	US-09-917-800A-1575	Sequence 1575, Ap
	32	26.4	26.4	2562	9	US-09-147-947-5	Sequence 5, Appli
	33	26.4	26.4	2614	9	US-09-147-947-3	Sequence 3, Appli
	34	26.2	26.2	1839	10	US-09-950-294-3	Sequence 3, Appli
	35	26.2	26.2	1845	9	US-10-047-542-12	Sequence 12, Appl
	36	26.2	26.2	2919	9	US-09-982-107-3	Sequence 3, Appli
	37	26.2	26.2	2919	9	US-09-989-919-74	Sequence 74, Appl
	38	26.2	26.2	3149	9	US-09-981-353-58	Sequence 58, Appl
	39	26.2	26.2	7129	9	US-10-047-542-101	Sequence 101, App
	40	26.2	26.2	8074	9	US-10-047-542-14	Sequence 14, Appl
c	41	26	26.0	681	9	US-10-078-770-91	Sequence 91, Appl
	42	25.8	25.8	2063	10	US-09-925-302-169	Sequence 169, App
	43	25.8	25.8	2386	9	US-09-808-602-92	Sequence 92, Appl
	44	25.8	25.8	2493	10	US-09-880-107-2256	Sequence 2256, Ap
	45	25.8	25.8	2555	12	US-10-044-090-613	Sequence 613, App

ALIGNMENTS

RESULT 1
US-09-808-602-100
; Sequence 100, Application US/09808602
; Patent No. US20020155115A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard A
; APPLICANT: Herrman, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishnu
; APPLICANT: Mezes, Peter S
; APPLICANT: MacDougall, John
; TITLE OF INVENTION: No. US20020155115Alel Proteins and Nuclec Acids Encoding Same
; FILE REFERENCE: 15966-697 CIP
; CURRENT APPLICATION NUMBER: US/09/808,602
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 09/800,198
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 100
; LENGTH: 1047
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-808-602-100

Query Match 36.0%; Score 36; DB 9; Length 1047;
Best Local Similarity 64.3%; Pred. No. 0.00054;
Matches 54; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

```
Qy      17 TGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTCTGAGCGAGTTCTACA 76
           ||  |||  | ||     || |   |||||| || ||  ||||||| || ||||  |  |
Db     207 TGAGCATTCAGGAAAGTGACCACCATTTCTGCGGCGGCTCCATTCTCAGTGAGTGGTGGA 266

Qy      77 TCCTAACGGCAGCCCACTGTCTCT 100
           |||| || |  ||||||||  |||
Db     267 TCCTCACCGTGGCCCACTGCTTCT 290
```

```
RESULT 2
US-09-764-898-72
; Sequence 72, Application US/09764898
; Patent No. US20020090673A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PJZ01
;  CURRENT APPLICATION NUMBER: US/09/764,898
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - consult PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 311
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 72
;   LENGTH: 1795
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-764-898-72

  Query Match            34.2%;  Score 34.2;  DB 10;  Length 1795;
  Best Local Similarity  64.6%;  Pred. No. 0.0032;
  Matches   51;  Conservative    0;  Mismatches   28;  Indels    0;  Gaps    0;

Qy      21 CATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTCTGAGCGAGTTCTACATCCT 80
           |||| |  | || ||  ||    ||||||| || | ||| |||  |     ||   ||||
Db     840 CATCCAGTACGACAAACAGCACGTCTGTGGAGGGAGCATCCTGGACCCCCACTGGGTCCT 899

Qy      81 AACGGCAGCCCACTGTCTC 99
           |||||||||||||||  ||
Db     900 CACGGCAGCCCACTGCTTC 918


RESULT 3
US-10-063-547-111
; Sequence 111, Application US/10063547
; Publication No. US20020182638A1
; GENERAL INFORMATION:
;  APPLICANT: Eaton,Dan L.
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Grimaldi,Christopher J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Watanabe,Colin K.
;  APPLICANT:  Wood,William I.
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
   FILE REFERENCE: P3230R1C1
   CURRENT APPLICATION NUMBER: US/10/063,547
   CURRENT FILING DATE:  2002-05-02
;  Prior Application removed - See File Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 111
;   LENGTH: 2063
;   TYPE: DNA
;   ORGANISM: Homo Sapien
US-10-063-547-111

  Query Match            34.2%;  Score 34.2;  DB 9;  Length 2063;
  Best Local Similarity  64.6%;  Pred. No. 0.0034;
  Matches   51;  Conservative    0;  Mismatches   28;  Indels    0;  Gaps    0;

Qy      21 CATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCAT CTGAGCGAGTTCTACATCCT 80
           |||| |  | || ||  ||    ||||||| || | ||| |||  |     ||   ||||
Db     857 CATCCAGTACGACAAACAGCACGTCTGTGGAGGGAGCATCCTGGACCCCCACTGGGTCCT 916

Qy      81 AACGGCAGCCCACTGTCTC 99
           |||||||||||||||  ||
Db     917 CACGGCAGCCCACTGCTTC 935


RESULT 4
US-10-174-590-329
; Sequence 329, Application US/10174590
; Publication No. US20030008352A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Chen,Jian
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Pan,James
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Watanabe,Colin K.
;  APPLICANT:  Wood,William I.
;  APPLICANT:  Zhang,Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3430R1C42
;  CURRENT APPLICATION NUMBER: US/10/174,590
;  CURRENT FILING DATE:  2002-06-18
;  Prior application removed - See File Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 329
;   LENGTH: 2063
;   TYPE: DNA
;   ORGANISM: Homo Sapien
US-10-174-590-329

  Query Match            34.2%;  Score 34.2;  DB 9;  Length 2063;
  Best Local Similarity  64.6%;  Pred. No. 0.0034;
  Matches   51;  Conservative    0;  Mismatches   28;  Indels    0;  Gaps    0;

Qy      21 CATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTCTGAGCGAGTTCTACATCCT 80
           |||| |  | || ||  ||    ||||||| || | ||| |||  |     ||   ||||
Db     857 CATCCAGTACGACAAACAGCACGTCTGTGGAGGGAGCATCCTGGACCCCCACTGGGTCCT 916

Qy      81 AACGGCAGCCCACTGTCTC 99
           |||||||||||||||  ||
Db     917 CACGGCAGCCCACTGCTTC 935


RESULT 5
US-10-176-758-329
; Sequence 329, Application US/10176758
; Publication No. US20030008353A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Chen,Jian
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Pan,James
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Watanabe,Colin K.
;  APPLICANT:  Wood,William I.
;  APPLICANT:  Zhang,Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3430R1C104
;  CURRENT APPLICATION NUMBER: US/10/176,758
;  CURRENT FILING DATE:  2002-06-21
;  Prior Application removed - See File Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 329
;   LENGTH: 2063
;   TYPE: DNA
;   ORGANISM: Homo Sapien
US-10-176-758-329

  Query Match            34.2%;  Score 34.2;  DB 9;  Length 2063;
  Best Local Similarity  64.6%;  Pred. No. 0.0034;
  Matches   51;  Conservative    0;  Mismatches   28;  Indels    0;  Gaps    0;
```

```
Qy      21 CATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTCTGAGCGAGTTCTACATCCT 80
           |||| |  | || ||  ||    ||||||| || | ||| |||  |     ||   ||||
Db     857 CATCCAGTACGACAAACAGCACGTCTGTGGAGGGAGCATCCTGGACCCCCACTGGGTCCT 916

Qy      81 AACGGCAGCCCACTGTCTC 99
           |||||||||||||||  ||
Db     917 CACGGCAGCCCACTGCTTC 935


RESULT 6
US-10-006-867-111
; Sequence 111, Application US/10006867
; Patent No. US20020119130A1
; GENERAL INFORMATION:
;  APPLICANT: Eaton,Dan L.
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Grimaldi,Christopher J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Watanabe,Colin K.
;  APPLICANT:  Wood,William I.
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3230R1C1
;  CURRENT APPLICATION NUMBER: US/10/006,867
;  CURRENT FILING DATE:  2001-12-06
;  PRIOR APPLICATION NUMBER: 60/063435
;  PRIOR FILING DATE: 1997-10-29
;  PRIOR APPLICATION NUMBER: 60/064215
;  PRIOR FILING DATE: 1997-10-29
;  PRIOR APPLICATION NUMBER: 60/082797
;  PRIOR FILING DATE: 1998-04-22
;  PRIOR APPLICATION NUMBER: 60/083495
;  PRIOR FILING DATE: 1998-04-29
;  PRIOR APPLICATION NUMBER: 60/085579
;  PRIOR FILING DATE: 1998-05-15
;  PRIOR APPLICATION NUMBER: 60/087759
;  PRIOR FILING DATE: 1998-06-02
;  PRIOR APPLICATION NUMBER: 60/088021
;  PRIOR FILING DATE: 1998-06-04
;  PRIOR APPLICATION NUMBER: 60/088029
;  PRIOR FILING DATE: 1998-06-04
;  PRIOR APPLICATION NUMBER: 60/088030
;  PRIOR FILING DATE: 1998-06-04
;  PRIOR APPLICATION NUMBER: 60/088734
;  PRIOR FILING DATE: 1998-06-10
;  PRIOR APPLICATION NUMBER: 60/088740
;  PRIOR FILING DATE: 1998-06-10
;  PRIOR APPLICATION NUMBER: 60/088811
;  PRIOR FILING DATE: 1998-06-10
;  PRIOR APPLICATION NUMBER: 60/088824
;  PRIOR FILING DATE: 1998-06-10
;  PRIOR APPLICATION NUMBER: 60/088825
;  PRIOR FILING DATE: 1998-06-10
;  PRIOR APPLICATION NUMBER: 60/088863
;  PRIOR FILING DATE: 1998-06-11
;  PRIOR APPLICATION NUMBER: 60/089105
;  PRIOR FILING DATE: 1998-06-12
;  PRIOR APPLICATION NUMBER: 60/089514
;  PRIOR FILING DATE: 1998-06-16
;  PRIOR APPLICATION NUMBER: 60/089653
;  PRIOR FILING DATE: 1998-06-17
;  PRIOR APPLICATION NUMBER: 60/089952
;  PRIOR FILING DATE: 1998-06-19
;  PRIOR APPLICATION NUMBER: 60/090246
;  PRIOR FILING DATE: 1998-06-22
;  PRIOR APPLICATION NUMBER: 60/090444
;  PRIOR FILING DATE: 1998-06-24
;  PRIOR APPLICATION NUMBER: 60/090688
;  PRIOR FILING DATE: 1998-06-25
;  PRIOR APPLICATION NUMBER: 60/090696
;  PRIOR FILING DATE: 1998-06-25
;  PRIOR APPLICATION NUMBER: 60/090862
;  PRIOR FILING DATE: 1998-06-26
;  PRIOR APPLICATION NUMBER: 60/091628
;  PRIOR FILING DATE: 1998-07-02
;  PRIOR APPLICATION NUMBER: 60/096012
;  PRIOR FILING DATE: 1998-08-10
;  PRIOR APPLICATION NUMBER: 60/096757
;  PRIOR FILING DATE: 1998-08-17
;  PRIOR APPLICATION NUMBER: 60/096949
;  PRIOR FILING DATE: 1998-08-18
;  PRIOR APPLICATION NUMBER: 60/096959
;  PRIOR FILING DATE: 1998-08-18
;  PRIOR APPLICATION NUMBER: 60/097954
;  PRIOR FILING DATE: 1998-08-26
;  PRIOR APPLICATION NUMBER: 60/097971
;  PRIOR FILING DATE: 1998-08-26
;  PRIOR APPLICATION NUMBER: 60/097979
;  PRIOR FILING DATE: 1998-08-26
;  PRIOR APPLICATION NUMBER: 60/098749
;  PRIOR FILING DATE: 1998-09-01
;  PRIOR APPLICATION NUMBER: 60/099741
;  PRIOR FILING DATE: 1998-09-10
;  PRIOR APPLICATION NUMBER: 60/099763
;  PRIOR FILING DATE: 1998-09-10
;  PRIOR APPLICATION NUMBER: 60/099792
;  PRIOR FILING DATE: 1998-09-10
;  PRIOR APPLICATION NUMBER: 60/099812
;  PRIOR FILING DATE: 1998-09-10
;  PRIOR APPLICATION NUMBER: 60/099815
;  PRIOR FILING DATE: 1998-09-10
;  PRIOR APPLICATION NUMBER: 60/100627
;  PRIOR FILING DATE: 1998-09-16
;  PRIOR APPLICATION NUMBER: 60/100662
;  PRIOR FILING DATE: 1998-09-16
;  PRIOR APPLICATION NUMBER: 60/100683
;  PRIOR FILING DATE: 1998-09-17
;  PRIOR APPLICATION NUMBER: 60/100684
;  PRIOR FILING DATE: 1998-09-17
;  PRIOR APPLICATION NUMBER: 60/100930
;  PRIOR FILING DATE: 1998-09-17
;  PRIOR APPLICATION NUMBER: 60/101279
;  PRIOR FILING DATE: 1998-09-22
;  PRIOR APPLICATION NUMBER: 60/101475
;  PRIOR FILING DATE: 1998-09-23
;  PRIOR APPLICATION NUMBER: 60/101738
;  PRIOR FILING DATE: 1998-09-24
;  PRIOR APPLICATION NUMBER: 60/101743
;  PRIOR FILING DATE: 1998-09-24
;  PRIOR APPLICATION NUMBER: 60/101916
;  PRIOR FILING DATE: 1998-09-24
;  PRIOR APPLICATION NUMBER: 60/102570
;  PRIOR FILING DATE: 1998-09-30
;  PRIOR APPLICATION NUMBER: 60/103449
;  PRIOR FILING DATE: 1998-10-06
;  PRIOR APPLICATION NUMBER: 60/103678
;  PRIOR FILING DATE: 1998-10-08
;  PRIOR APPLICATION NUMBER: 60/103679
;  PRIOR FILING DATE: 1998-10-08
;  PRIOR APPLICATION NUMBER: 60/103711
;  PRIOR FILING DATE: 1998-10-08
;  PRIOR APPLICATION NUMBER: 60/105000
;  PRIOR FILING DATE: 1998-10-20
;  PRIOR APPLICATION NUMBER: 60/105002
;  PRIOR FILING DATE: 1998-10-20
;  PRIOR APPLICATION NUMBER: 60/105881
;  PRIOR FILING DATE: 1998-10-27
;  PRIOR APPLICATION NUMBER: 60/106030
;  PRIOR FILING DATE: 1998-10-28
;  PRIOR APPLICATION NUMBER: 60/106464
;  PRIOR FILING DATE: 1998-10-30
;  PRIOR APPLICATION NUMBER: 60/106856
```

```
;  PRIOR FILING DATE: 1998-11-03
;  PRIOR APPLICATION NUMBER: 60/108807
;  PRIOR FILING DATE: 1998-11-17
;  PRIOR APPLICATION NUMBER: 60/112419
;  PRIOR FILING DATE: 1998-12-15
;  PRIOR APPLICATION NUMBER: 60/112422
;  PRIOR FILING DATE: 1998-12-15
;  PRIOR APPLICATION NUMBER: 60/112853
;  PRIOR FILING DATE: 1998-12-16
;  PRIOR APPLICATION NUMBER: 60/113011
;  PRIOR FILING DATE: 1998-12-16
;  PRIOR APPLICATION NUMBER: 60/112854
;  PRIOR FILING DATE: 1998-12-16
;  PRIOR APPLICATION NUMBER: 60/113300
;  PRIOR FILING DATE: 1998-12-22
;  PRIOR APPLICATION NUMBER: 60/113408
;  PRIOR FILING DATE: 1998-12-22
;  PRIOR APPLICATION NUMBER: 60/113430
;  PRIOR FILING DATE: 1998-12-23
;  PRIOR APPLICATION NUMBER: 60/113621
;  PRIOR FILING DATE: 1998-12-23
;  PRIOR APPLICATION NUMBER: 60/114223
;  PRIOR FILING DATE: 1998-12-30
;  PRIOR APPLICATION NUMBER: 60/115614
;  PRIOR FILING DATE: 1999-01-12
;  PRIOR APPLICATION NUMBER: 60/116527
;  PRIOR FILING DATE: 1999-01-20
;  PRIOR APPLICATION NUMBER: 60/116843
;  PRIOR FILING DATE: 1999-01-22
;  PRIOR APPLICATION NUMBER: 60/119285
;  PRIOR FILING DATE: 1999-02-09
;  PRIOR APPLICATION NUMBER: 60/119287
;  PRIOR FILING DATE: 1999-02-09
;  PRIOR APPLICATION NUMBER: 60/119525
;  PRIOR FILING DATE: 1999-02-10
;  PRIOR APPLICATION NUMBER: 60/119549
;  PRIOR FILING DATE: 1999-02-10
;  PRIOR APPLICATION NUMBER: 60/120014
;  PRIOR FILING DATE: 1999-02-11
;  PRIOR APPLICATION NUMBER: 60/129122
;  PRIOR FILING DATE: 1999-04-13
;  PRIOR APPLICATION NUMBER: 60/129674
;  PRIOR FILING DATE: 1999-04-16
;  PRIOR APPLICATION NUMBER: 60/131291
;  PRIOR FILING DATE: 1999-04-27
;  PRIOR APPLICATION NUMBER: 60/138387
;  PRIOR FILING DATE: 1999-06-09
;  PRIOR APPLICATION NUMBER: 60/144791
;  PRIOR FILING DATE: 1999-07-20
;  PRIOR APPLICATION NUMBER: 60/169495
;  PRIOR FILING DATE: 1999-12-07
;  PRIOR APPLICATION NUMBER: 60/175481
;  PRIOR FILING DATE: 2000-01-11
;  PRIOR APPLICATION NUMBER: 60/191007
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/199397
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: 09/380139
;  PRIOR FILING DATE: 1998-08-25
;  PRIOR APPLICATION NUMBER: 09/311832
;  PRIOR FILING DATE: 1999-05-14
;  PRIOR APPLICATION NUMBER: 09/380137
;  PRIOR FILING DATE: 1999-08-25
;  PRIOR APPLICATION NUMBER: 09/380138
;  PRIOR FILING DATE: 1999-08-25
;  PRIOR APPLICATION NUMBER: 09/380142

  Query Match           34.2%;  Score 34.2;  DB 12;  Length 2063;
  Best Local Similarity  64.6%;  Pred. No. 0.0034;
  Matches   51;  Conservative    0;  Mismatches   28;  Indels    0;  Gaps    0;

Qy      21 CATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTCTGAGCGAGTTCTACATCCT 80
           |||| |  | || ||  ||    ||||||| || | ||| |||  |     ||   ||||

Db     857 CATCCAGTACGACAAACAGCACGTCTGTGGAGGGAGCATCCTGGACCCCCACTGGGTCCT 916

Qy      81 AACGGCAGCCCACTGTCTC 99
           ||||||||||||||  ||
Db     917 CACGGCAGCCCACTGCTTC 935


RESULT 7
US-10-052-586-329
; Sequence 329, Application US/10052586
; Patent No. US20020127584A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Chen,Jian
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Pan,James
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Watanabe,Colin K.
;  APPLICANT:  Wood,William I.
;  APPLICANT:  Zhang,Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3430R1C1
;  CURRENT APPLICATION NUMBER: US/10/052,586
;  CURRENT FILING DATE:  2002-01-15
;  PRIOR APPLICATION NUMBER: 60/059263
;  PRIOR FILING DATE: 1997-09-18
;  PRIOR APPLICATION NUMBER: 60/059266
;  PRIOR FILING DATE: 1997-09-18
;  PRIOR APPLICATION NUMBER: 60/062250
;  PRIOR FILING DATE: 1997-10-17
;  PRIOR APPLICATION NUMBER: 60/063120
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: 60/063121
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: 60/063486
;  PRIOR FILING DATE: 1997-10-21
;  PRIOR APPLICATION NUMBER: 60/063540
;  PRIOR FILING DATE: 1997-10-28
;  PRIOR APPLICATION NUMBER: 60/063541
;  PRIOR FILING DATE: 1997-10-28
;  PRIOR APPLICATION NUMBER: 60/063544
;  PRIOR FILING DATE: 1997-10-28
;  PRIOR APPLICATION NUMBER: 60/063564
;  PRIOR FILING DATE: 1997-10-28
;  PRIOR APPLICATION NUMBER: 60/063734
;  PRIOR FILING DATE: 1997-10-29
;  PRIOR APPLICATION NUMBER: 60/063870
;  PRIOR FILING DATE: 1997-10-31
;  PRIOR APPLICATION NUMBER: 60/064103
;  PRIOR FILING DATE: 1997-10-31
;  PRIOR APPLICATION NUMBER: 60/065311
;  PRIOR FILING DATE: 1997-11-13
;  PRIOR APPLICATION NUMBER: 60/066120
;  PRIOR FILING DATE: 1997-11-21
;  PRIOR APPLICATION NUMBER: 60/066466
;  PRIOR FILING DATE: 1997-11-24
;  PRIOR APPLICATION NUMBER: 60/066772
;  PRIOR FILING DATE: 1997-11-24
;  PRIOR APPLICATION NUMBER: 60/069335
;  PRIOR FILING DATE: 1997-12-11
;  PRIOR APPLICATION NUMBER: 60/069425
;  PRIOR FILING DATE: 1997-12-12
;  PRIOR APPLICATION NUMBER: 60/069870
;  PRIOR FILING DATE: 1997-12-17
;  PRIOR APPLICATION NUMBER: 60/068017
;  PRIOR FILING DATE: 1997-12-18
;  PRIOR APPLICATION NUMBER: 60/077450
;  PRIOR FILING DATE: 1998-03-10
;  PRIOR APPLICATION NUMBER: 60/077632
```

; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/078886
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079786
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/080107
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080194
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080327
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081070
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081195
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081838
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082568
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082569
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082704
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082797
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083496
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083499
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/086023
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/086392
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/086486
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087098
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087208
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088722
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088740
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088811
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088825
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088863
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089090
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089908

Query Match 34.2%; Score 34.2; DB 12; Length 2063;
Best Local Similarity 64.6%; Pred. No. 0.0034;
Matches 51; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

```
Qy      21 CATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTCTGAGCGAGTTCTACATCCT 80
           |||| |  | || ||  ||    ||||||| || | ||| |||  |     ||   ||||
Db     857 CATCCAGTACGACAAACAGCACGTCTGTGGAGGGAGCATCCTGGACCCCCACTGGGTCCT 916

Qy      81 AACGGCAGCCCACTGTCTC 99
```

```
          ||||||||||||||  ||
Db    917 CACGGCAGCCCACTGCTTC 935


RESULT 8
US-09-851-588-5
; Sequence 5, Application US/09851588
; Patent No. US20020042067A1
; GENERAL INFORMATION:
;  APPLICANT: Mack, David
;  APPLICANT:  Gish, Kurt C.
;  APPLICANT:  Wilson, Keith E.
;  TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSING COLORECTAL CANCER, COMPOSITIONS, AND
;  TITLE OF INVENTION:  OF SCREENING FOR COLORECTAL CANCER MODULATORS
;  FILE REFERENCE: A-68829-1/DJB/JJD/AMS
;  CURRENT APPLICATION NUMBER: US/09/851,588
;  CURRENT FILING DATE:  2001-09-24
;  PRIOR APPLICATION NUMBER: US 09/642,252
;  PRIOR FILING DATE: 2000-08-17
;  PRIOR APPLICATION NUMBER: US 09/656,002
;  PRIOR FILING DATE: 2000-09-06
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
;   LENGTH: 2079
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-851-588-5

  Query Match             34.2%;  Score 34.2;  DB 10;  Length 2079;
  Best Local Similarity   64.6%;  Pred. No. 0.0034;
  Matches   51;  Conservative    0;  Mismatches   28;  Indels    0;  Gaps    0;

Qy      21 CATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTCTGAGCGAGTTCTACATCCT 80
           ||||  |  | ||  ||  ||    ||||||| || | ||| |||  |     ||   ||||
Db     871 CATCCAGTACGACAAACAGCACGTCTGTGGAGGGAGCATCCTGGACCCCCACTGGGTCCT 930

Qy      81 AACGGCAGCCCACTGTCTC 99
           |||||||||||||||  ||
Db     931 CACGGCAGCCCACTGCTTC 949


RESULT 9
US-09-851-588-7
; Sequence 7, Application US/09851588
; Patent No. US20020042067A1
; GENERAL INFORMATION:
;  APPLICANT: Mack, David
;  APPLICANT:  Gish, Kurt C.
;  APPLICANT:  Wilson, Keith E.
;  TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSING COLORECTAL CANCER, COMPOSITIONS, AND
;  TITLE OF INVENTION:  OF SCREENING FOR COLORECTAL CANCER MODULATORS
;  FILE REFERENCE: A-68829-1/DJB/JJD/AMS
;  CURRENT APPLICATION NUMBER: US/09/851,588
;  CURRENT FILING DATE:  2001-09-24
;  PRIOR APPLICATION NUMBER: US 09/642,252
;  PRIOR FILING DATE: 2000-08-17
;  PRIOR APPLICATION NUMBER: US 09/656,002
;  PRIOR FILING DATE: 2000-09-06
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
;   LENGTH: 2081
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (215)..(1528)
;   OTHER INFORMATION:
US-09-851-588-7

  Query Match             34.2%;  Score 34.2;  DB 10;  Length 2081;
  Best Local Similarity   64.6%;  Pred. No. 0.0034;
  Matches   51;  Conservative    0;  Mismatches   28;  Indels    0;  Gaps    0;

Qy      21 CATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTCTGAGCGAGTTCTACATCCT 80
           ||||  |  | ||  ||  ||    ||||||| || | ||| |||  |     ||   ||||
Db     877 CATCCAGTACGACAAACAGCACGTCTGTGGAGGGAGCATCCTGGACCCCCACTGGGTCCT 936

Qy      81 AACGGCAGCCCACTGTCTC 99
           |||||||||||||||  ||
Db     937 CACGGCAGCCCACTGCTTC 955


RESULT 10
US-09-960-352-10052
; Sequence 10052, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
;  APPLICANT: Warren, Wesley C.
;  APPLICANT:  Tao, Nengbing
;  APPLICANT:  Byatt, John C.
;  APPLICANT:  Mathialagan, Nagappan
;  TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AN
;  TITLE OF INVENTION:  MUSCLE AND FAT DEPOSITION
;  FILE REFERENCE: 16511.006/37-21(10298)C
;  CURRENT APPLICATION NUMBER: US/09/960,352
;  CURRENT FILING DATE:  2001-09-24
;  NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 10052
;   LENGTH: 396
;   TYPE: DNA
;   ORGANISM: Bos taurus
;   OTHER INFORMATION: Clone ID: 43-LIB34-017-Q1-E1-C4
US-09-960-352-10052

  Query Match             30.2%;  Score 30.2;  DB 10;  Length 396;
  Best Local Similarity   60.2%;  Pred. No. 0.056;
  Matches   50;  Conservative    0;  Mismatches   33;  Indels    0;  Gaps    0;

Qy      17 TGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTCTGAGCGAGTTCTACA 76
           | ||  |  |||  ||||  |  |  |||||||| ||  ||||  | |  ||    |
Db      60 TCCTTTTGCATGGTGAAATTGCTGCATTCTGTGGAGGTTCCATCGTTAATGAAAAATGGG 119

Qy      77 TCCTAACGGCAGCCCACTGTCTC 99
           |  |||| ||||||||||  ||
Db     120 TTGTAACTGCAGCCCACTGCATC 142


RESULT 11
US-09-960-352-6361
; Sequence 6361, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
;  APPLICANT: Warren, Wesley C.
;  APPLICANT:  Tao, Nengbing
;  APPLICANT:  Byatt, John C.
;  APPLICANT:  Mathialagan, Nagappan
;  TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AN
;  TITLE OF INVENTION:  MUSCLE AND FAT DEPOSITION
;  FILE REFERENCE: 16511.006/37-21(10298)C
;  CURRENT APPLICATION NUMBER: US/09/960,352
;  CURRENT FILING DATE:  2001-09-24
;  NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 6361
;   LENGTH: 420
;   TYPE: DNA
;   ORGANISM: Bos taurus
;   OTHER INFORMATION: Clone ID: 27-LIB34-043-Q1-E1-G3
US-09-960-352-6361

  Query Match             30.2%;  Score 30.2;  DB 10;  Length 420;
  Best Local Similarity   60.2%;  Pred. No. 0.057;
  Matches   50;  Conservative    0;  Mismatches   33;  Indels    0;  Gaps    0;

Qy      17 TGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTCTGAGCGAGTTCTACA 76
```

```
          | ||  |  |||  ||||  |  |  |||||||| ||  ||||  | |  ||    |
Db      29 TCCTTTTGCATGGTGAAATTGCTGCATTCTGTGGAGGTTCCATCGTTAATGAAAAATGGG 88

Qy      77 TCCTAACGGCAGCCCACTGTCTC 99
           |  |||| ||||||||||| ||
Db      89 TTGTAACTGCAGCCCACTGCATC 111


RESULT 12
US-09-738-626-1340/c
; Sequence 1340, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAGAWA, SATOSHI
;  APPLICANT:  MIZOGUCHI, HIROSHI
;  APPLICANT:  ANDO, SEIKO
;  APPLICANT:  HAYASHI, MIKIRO
;  APPLICANT:  OCHIAI, KEIKO
;  APPLICANT:  YOKOI, HARUHIKO
;  APPLICANT:  TATEISHI, NAOKO
;  APPLICANT:  SENOH, AKIHIRO
;  APPLICANT:  IKEDA, MASATO
;  APPLICANT:  OZAKI, AKIO
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;  FILE REFERENCE: 249-125
;  CURRENT APPLICATION NUMBER: US/09/738,626
;  CURRENT FILING DATE:  2000-12-18
;  PRIOR APPLICATION NUMBER: JP 99/377484
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: JP 00/159162
;  PRIOR FILING DATE: 2000-04-07
;  PRIOR APPLICATION NUMBER: JP 00/280988
;  PRIOR FILING DATE: 2000-08-03
;  NUMBER OF SEQ ID NOS: 7059
;  SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 1340
;   LENGTH: 1674
;   TYPE: DNA
;   ORGANISM: Corynebacterium glutamicum
US-09-738-626-1340

  Query Match             29.8%;  Score 29.8;  DB 9;  Length 1674;
  Best Local Similarity   66.2%;  Pred. No. 0.14;
  Matches   43;  Conservative    0;  Mismatches   22;  Indels    0;  Gaps    0;

Qy      13 GCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTCTGAGCGAGTTC 72
           |||||||||  || | |      |   |||| | |  ||||||| || ||||||| ||| | |
Db     714 GCCCTGCTCCTCGAGGGTCTTACGCAGGGCTGCAATGGTGGAGCCCTTCTGACCGATTGC 655

Qy      73 TACAT 77
           || |
Db     654 GACGT 650


RESULT 13
US-09-888-615-36
; Sequence 36, Application US/09888615
; Patent No. US20020064856A1
; GENERAL INFORMATION:
;  APPLICANT: PLOWMAN, GREGORY
;  APPLICANT:  WHYTE, DAVID
;  APPLICANT:  CAENEPEEL, SEAN
;  APPLICANT:  CHARYDCZAK, GLEN
;  APPLICANT:  MANNING, GERARD
;  APPLICANT:  SUDARSANAM, SUCHA
;  TITLE OF INVENTION: NOVEL PROTEASES
;  FILE REFERENCE: 038602/1214
;  CURRENT APPLICATION NUMBER: US/09/888,615
;  CURRENT FILING DATE:  2001-06-26
;  PRIOR APPLICATION NUMBER: 60/214,047
;  PRIOR FILING DATE: 2000-06-26
;  NUMBER OF SEQ ID NOS: 150
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 36
;   LENGTH: 1059
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-888-615-36

  Query Match             29.6%;  Score 29.6;  DB 10;  Length 1059;
  Best Local Similarity   64.7%;  Pred. No. 0.14;
  Matches   44;  Conservative    0;  Mismatches   24;  Indels    0;  Gaps    0;

Qy      33 AAACGAGGGTTTCTGTGGTGGAACCATTCTGAGCGAGTTCTACATCCTAACGGCAGCCCA 92
           ||  ||   ||||||||| ||  |||| || | | |||  |  || || || || || ||
Db     264 AAGTGAACCTTTCTGTGGCGGCTCCATCCTCAACAAGTGGTGGATTCTCACTGCGGCTCA 323

Qy      93 CTGTCTCT 100
           |||  | |
Db     324 CTGCTTAT 331


RESULT 14
US-09-755-016-5
; Sequence 5, Application US/09755016
; Patent No. US20010034437A1
; GENERAL INFORMATION:
;  APPLICANT: Walke, D.  Wade
;  APPLICANT:  Wilganowski, Nathaniel L.
;  APPLICANT:  Donoho, Gregory
;  APPLICANT:  Turner, C. Alexander Jr.
;  TITLE OF INVENTION: No. US20010034437A1el Human Proteases and Polynucleotides Enc
;  FILE REFERENCE: LEX-0114-USA
;  CURRENT APPLICATION NUMBER: US/09/755,016
;  CURRENT FILING DATE:  2001-06-01
;  PRIOR APPLICATION NUMBER: US 60/174,686
;  PRIOR FILING DATE: 2000-01-06
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
;   LENGTH: 867
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-755-016-5

  Query Match             28.4%;  Score 28.4;  DB 10;  Length 867;
  Best Local Similarity   70.4%;  Pred. No. 0.36;
  Matches   38;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy      42 TTTCTGTGGTGGAACCATTCTGAGCGAGTTCTACATCCTAACGGCAGCCCACTG 95
           | | ||||| ||||||  |  |||| |||    |   |||| || || ||||||||
Db     135 TGTATGTGGGGGAACCCTAGTGAGAGAGAGGTGGGTCCTCACAGCTGCCCACTG 188


RESULT 15
US-09-755-016-3
; Sequence 3, Application US/09755016
; Patent No. US20010034437A1
; GENERAL INFORMATION:
;  APPLICANT: Walke, D.  Wade
;  APPLICANT:  Wilganowski, Nathaniel L.
;  APPLICANT:  Donoho, Gregory
;  APPLICANT:  Turner, C. Alexander Jr.
;  TITLE OF INVENTION: No. US20010034437A1el Human Proteases and Polynucleotides Enc
;  FILE REFERENCE: LEX-0114-USA
;  CURRENT APPLICATION NUMBER: US/09/755,016
;  CURRENT FILING DATE:  2001-06-01
;  PRIOR APPLICATION NUMBER: US 60/174,686
;  PRIOR FILING DATE: 2000-01-06
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 1047
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-755-016-3
```

Query Match 28.4%; Score 28.4; DB 10; Length 1047;
Best Local Similarity 70.4%; Pred. No. 0.39;
Matches 38; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

```
Qy      42 TTTCTGTGGTGGAACCATTCTGAGCGAGTTCTACATCCTAACGGCAGCCCACTG 95
           | | ||||| |||||| |  |||| |||   |   |||| || || ||||||||
Db     315 TGTATGTGGGGGAACCCTAGTGAGAGAGAGGTGGGTCCTCACAGCTGCCCACTG 368
```

Search completed: January 15, 2003, 22:32:23
Job time : 34.5 secs

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: January 15, 2003, 19:39:40 ; Search time 1268 Seconds
(without alignments)
1277.247 Million cell updates/sec

Title: L00396_COPY_1_100
Perfect score: 100
Sequence: 1 CGTCTGTCACAGGCCCTGCT.....AACGGCAGCCCACTGTCTCT 100

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	92.6	92.6	855	9	AL570367	AL570367 AL570367
c	2	92.2	92.2	806	9	AL576464	AL576464 AL576464
	3	91	91.0	246	12	BE766639	BE766639 IL3-NT010
	4	91	91.0	329	14	R02602	R02602 ye76a02.r1
	5	89.4	89.4	649	10	AV694819	AV694819 AV694819
	6	88.4	88.4	1049	13	BM546790	BM546790 AGENCOURT
	7	87.4	87.4	479	14	N77287	N77287 yv43c02.r1
	8	81.6	81.6	984	9	AL570383	AL570383 AL570383
c	9	80	80.0	820	9	AL521983	AL521983 AL521983
	10	79	79.0	898	9	AL521984	AL521984 AL521984
	11	70.8	70.8	521	12	BG609107	BG609107 322632 MA
	12	64.4	64.4	609	14	W70576	W70576 me42a04.r1
	13	64.4	64.4	617	14	W70579	W70579 me42b03.r1
	14	64.4	64.4	824	13	BI144826	BI144826 602910055
	15	58.8	58.8	582	9	AA237419	AA237419 mx19d10.r
	16	58.6	58.6	844	9	AL553470	AL553470 AL553470
	17	49	49.0	879	17	CNS02TO1	AL213418 Tetraodon
	18	49	49.0	1061	17	CNS039FM	AL233851 Tetraodon
c	19	44.6	44.6	768	9	AI192868	AI192868 qe68b06.x
	20	41.4	41.4	959	13	BM438664	BM438664 IpLvr0126
c	21	40.2	40.2	678	9	AA554502	AA554502 ni36e02.s
	22	38	38.0	565	12	BG641571	BG641571 pg11c.pk0
	23	38	38.0	598	12	BG641612	BG641612 pg11c.pk0
	24	38	38.0	640	12	BG641885	BG641885 pg11c.pk0
	25	37.2	37.2	474	9	AL835455	AL835455 AL835455
	26	36	36.0	169	9	AA140046	AA140046 mr73g12.r
	27	36	36.0	408	9	AI893715	AI893715 mr73h12.y
	28	36	36.0	783	9	AI526776	AI526776 uj42g03.y
	29	36	36.0	1047	11	AK016601	AK016601 Mus muscu
	30	35.8	35.8	445	14	BQ790628	BQ790628 92-M3-272
	31	34.8	34.8	618	13	BJ493141	BJ493141 BJ493141
	32	34.4	34.4	653	9	AI258617	AI258617 LP01850.5
	33	34.2	34.2	367	12	BF896867	BF896867 PM3-MT011
	34	34.2	34.2	460	9	AL694581	AL694581 AL694581
c	35	34.2	34.2	462	13	BG984172	BG984172 IL5-CN006
	36	34.2	34.2	498	9	AI158537	AI158537 ud25e02.r
	37	34.2	34.2	559	14	BM686292	BM686292 UI-E-CR0-
c	38	34.2	34.2	844	17	CNS04KQI	AL295155 Tetraodon
	39	34.2	34.2	920	14	BQ922656	BQ922656 AGENCOURT
	40	34.2	34.2	975	12	BG288427	BG288427 602388091
	41	34.2	34.2	1973	11	BC004855	BC004855 Homo sapi
	42	33.8	33.8	133	9	AA140049	AA140049 mr73h12.r
	43	33.8	33.8	944	17	CNS037K4	AL231421 Tetraodon
	44	33.6	33.6	522	13	BM536550	BM536550 ha71e11.g
	45	33.6	33.6	531	13	BM646546	BM646546 170006873

ALIGNMENTS

RESULT 1
AL570367/c
LOCUS AL570367 855 bp mRNA linear EST 16-FEB-2001
DEFINITION AL570367 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0DI020YH13 3 prime, mRNA sequence.
ACCESSION AL570367
VERSION AL570367.1 GI:12926604
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 855)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES Location/Qualifiers
source 1. .855
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DI020YH13"
/clone_lib="LTI_NFL006_PL2"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end

enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"
BASE COUNT 129 a 263 c 250 g 210 t 3 others
ORIGIN

Query Match 92.6%; Score 92.6; DB 9; Length 855;
Best Local Similarity 96.0%; Pred. No. 3.7e-22;
Matches 95; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy       2 GTCTGTCACAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 61
           |||  | ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     672 GTCCCTGGCAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 613

Qy      62 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 100
           |||||||||||||||||||||||||||||||||||||||
       612 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 574
```

RESULT 2
AL576464/c
LOCUS AL576464 806 bp mRNA linear EST 16-FEB-2001
DEFINITION AL576464 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0DI076YA10 3 prime, mRNA sequence.
ACCESSION AL576464
VERSION AL576464.1 GI:12938633
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 806)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES Location/Qualifiers
source 1. .806
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DI076YA10"
/clone_lib="LTI_NFL006_PL2"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"
BASE COUNT 124 a 251 c 239 g 188 t 4 others
ORIGIN

Query Match 92.2%; Score 92.2; DB 9; Length 806;
Best Local Similarity 94.9%; Pred. No. 4.9e-22;
Matches 94; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```
Qy       2 GTCTGTCACAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 61
           |||  | ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     692 GTCCCTGGCAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 633

Qy      62 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 100
           |||||||||||||||||||||||||||||||||||||||:|||||
Db     632 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTKTCTCT 594
```

RESULT 3
BE766639
LOCUS BE766639 246 bp mRNA linear EST 19-SEP-2000
DEFINITION IL3-NT0105-200700-220-A03 NT0105 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE766639
VERSION BE766639.1 GI:10196563
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 246)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=IL3-NT0105-200700-220-A03&t3=2000-07-20&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 246.
FEATURES Location/Qualifiers
source 1. .246
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NT0105"
/dev_stage="Adult"
/note="Organ: nervous_tumor; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
BASE COUNT 59 a 70 c 73 g 44 t
ORIGIN

Query Match 91.0%; Score 91; DB 12; Length 246;
Best Local Similarity 94.9%; Pred. No. 6.8e-22;
Matches 94; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy       2 GTCTGTCACAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 61
           |||  | ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db     116 GTCCCTGGCAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACTATTC 175

Qy      62 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 100
           |||||||||||||||||||||||||||||||||||||||
Db     176 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 214
```

RESULT 4
R02602
LOCUS R02602 329 bp mRNA linear EST 31-MAR-1995
DEFINITION ye76a02.r1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone

```
            IMAGE:123626 5' similar to gb:M57285 COAGULATION FACTOR X PRECURSOR
            (HUMAN);, mRNA sequence.
ACCESSION   R02602
VERSION     R02602.1  GI:752338
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 329)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
            ,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
            Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
            ,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 814
            High quality sequence stops: 265 Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 814    Std Error: 0.00
            Seq primer: M13RP1
            High quality sequence stop: 265.
FEATURES             Location/Qualifiers
     source          1. .329
                     /organism="Homo sapiens"
                     /db_xref="GDB:476171"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:123626"
                     /clone_lib="Soares fetal liver spleen 1NFLS"
                     /sex="male"
                     /dev_stage="20 week-post conception fetus"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
                     with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
                     1st strand cDNA was primed with a Pac I - oligo(dT) primer
                     [5' AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Pac I and cloned into the Pac I
                     and Eco RI sites of the modified pT7T3 vector.  Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT       73 a     89 c    104 g     61 t      2 others
ORIGIN

  Query Match             91.0%;  Score 91;  DB 14;  Length 329;
  Best Local Similarity   94.9%;  Pred. No. 8.1e-22;
  Matches   94;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       2 GTCTGTCACAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 61
           |||  |  |||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db      20 GTCCCTGGCAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACTATTC 79

Qy      62 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 100
           |||||||||||!|||||||||||||||||||||||||||
Db      80 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 118


RESULT 5
AV694819
LOCUS       AV694819                649 bp    mRNA    linear   EST 16-JAN-2002
DEFINITION  AV694819 GKC Homo sapiens cDNA clone GKCEEE12 5', mRNA sequence.
ACCESSION   AV694819
VERSION     AV694819.1  GI:10296682
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 649)
  AUTHORS   Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X.,
            Xiao,H., Qu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W.,
            Shen,K., Lu,G., Fu,G., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X.,
            Hu,G., Gu,J., Chen,Z. and Han,Z.
  TITLE     Insight into hepatocellular carcinogenesis at transcriptome level
            by comparing gene expression profiles of hepatocellular carcinoma
            with those of corresponding noncancerous liver
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
  MEDLINE   21625106
COMMENT     Contact: Zeguang Han
            Chinese National Human Genome Center at Shanghai
            351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
            201203, P. R. China
            Tel: 86-21-50801919(ex.45)
            Fax: 86-21-50801[illegible]2
            Email: hanzg@chgc.sh.cn
            This clone is available at CHGC in Shanghai.
FEATURES             Location/Qualifiers
     source          1. .649
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="GKCEEE12"
                     /clone_lib="GKC"
                     /tissue_type="hepatocellular carcinoma"
                     /dev_stage="Adult"
                     /lab_host="SOLR"
                     /note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
                     XhoI"
BASE COUNT      142 a    192 c    205 g    110 t
ORIGIN

  Query Match             89.4%;  Score 89.4;  DB 10;  Length 649;
  Best Local Similarity   98.9%;  Pred. No. 4.4e-21;
  Matches   90;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      10 CAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTCTGAGCGAG 69
           ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db       6 CAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACTATTCTGAGCGAG 65

Qy      70 TTCTACATCCTAACGGCAGCCCACTGTCTCT 100
           |||||||||||||||||||||||||||||||
Db      66 TTCTACATCCTAACGGCAGCCCACTGTCTCT 96


RESULT 6
BM546790
LOCUS       BM546790                1049 bp    mRNA    linear   EST 20-FEB-2002
DEFINITION  AGENCOURT_6491143 NIH_MGC_125 Homo sapiens cDNA clone IMAGE:5723510
            5', mRNA sequence.
ACCESSION   BM546790
VERSION     BM546790.1  GI:18780039
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1049)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Invitrogen
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12711  row: e  column: 15
```

```
            High quality sequence stop: 759.
FEATURES             Location/Qualifiers
     source          1. .1049
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5723510"
                     /clone_lib="NIH_MGC_125"
                     /lab_host="DH10B"
                     /note="Organ: ovary (pool of 3); Vector: pCMV-SPORT6;
                     Site_1: EcoRV (destroyed); Site_2: NotI; RNA source pool
                     of three ovaries, from females ranging in age from 38 to
                     49 yo. Library is oligo-dT primed and directionally cloned
                     (EcoRV site is destroyed upon cloning). Average insert
                     size 2.1 kb, insert size range 1-3.5 kb. Library is
                     normalized and enriched for full-length clones and was
                     constructed by C. Gruber (Invitrogen). Research Genetics
                     tracking code 036."
BASE COUNT      274 a    289 c    316 g    169 t      1 others
ORIGIN

  Query Match             88.4%;  Score 88.4;  DB 13;  Length 1049;
  Best Local Similarity   92.9%;  Pred. No. 1.3e-20;
  Matches   92;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       2 GTCTGTCACAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 61
           |||   |  |||||||||||||||||||| |||||||||||||||||||||| || ||||
Db     795 GTCCCTGGCAGGCCCTGCTCATCAATGANGAAAACGAGGGTTTCTGTGGTGGGACTATTC 854

Qy      62 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 100
           |||||||||||||||||||||||||||||||||||||||
Db     855 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 893


RESULT 7
N77287
LOCUS       N77287                   479 bp    mRNA    linear   EST 28-JAN-1997
DEFINITION  yv43c02.r1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone
            IMAGE:245474 5' similar to gb:M57285 COAGULATION FACTOR X PRECURSOR
            (HUMAN);, mRNA sequence.
ACCESSION   N77287
VERSION     N77287.1  GI:1239865
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 479)
  AUTHORS   Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
            Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins
            ,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore
            ,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T.,
            Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J., Trevaskis,E.,
            Underwood,K., Wohldmann,P., Waterston,R., Wilson,R. and Marra,M.
  TITLE     Generation and analysis of 280,000 human expressed sequence tags
  JOURNAL   Genome Res. 6 (9), 807-828 (1996)
  MEDLINE   97044478
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 840   Std Error: 0.00
            Seq primer: reverse ET
            High quality sequence stop: 425.
FEATURES             Location/Qualifiers
     source          1. .479
                     /organism="Homo sapiens"
                     /db_xref="GDB:3794720"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:245474"
                     /clone_lib="Soares fetal liver spleen 1NFLS"
                     /sex="male"
                     /dev_stage="20 week-post conception fetus"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
                     with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
                     1st strand cDNA was primed with a Pac I - oligo(dT) primer
                     [5' AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Pac I and cloned into the Pac I
                     and Eco RI sites of the modified pT7T3 vector.  Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT      109 a    137 c    145 g     83 t      5 others
ORIGIN

  Query Match             87.4%;  Score 87.4;  DB 14;  Length 479;
  Best Local Similarity   97.8%;  Pred. No. 1.9e-20;
  Matches   88;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      11 AGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTCTGAGCGAGT 70
           |||||||| ||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db      10 AGGCCCTGNTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACTATTCTGAGCGAGT 69

Qy      71 TCTACATCCTAACGGCAGCCCACTGTCTCT 100
           ||||||||||||||||||||||||||||||
Db      70 TCTACATCCTAACGGCAGCCCACTGTCTCT 99


RESULT 8
AL570383
LOCUS       AL570383                 984 bp    mRNA    linear   EST 16-FEB-2001
DEFINITION  AL570383 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0DI020YH13 5
            prime, mRNA sequence.
ACCESSION   AL570383
VERSION     AL570383.1  GI:12926636
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 984)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             Location/Qualifiers
     source          1. .984
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CS0DI020YH13"
                     /clone_lib="LTI_NFL006_PL2"
                     /tissue_type="placenta"
                     /note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer.  Five prime end
                     enriched, double-stranded cDNA was digested with Not I and
                     cloned into the Not I and Eco RV sites of the pCMVSPORT 6
                     vector. Library was normalized. Library was constructed by
                     Life Technologies. Contact : Feng Liang Life Technologies,
                     a division of Invitrogen 9800 Medical Center Drive
                     Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
                     Email : fliang@lifetech.com URL :
                     http://fulllength.invitrogen.com"
BASE COUNT      252 a    269 c    300 g    160 t      3 others
ORIGIN

  Query Match             81.6%;  Score 81.6;  DB 9;  Length 984;
  Best Local Similarity   95.0%;  Pred. No. 3.4e-18;
  Matches   95;  Conservative    0;  Mismatches    4;  Indels    1;  Gaps    1;
```

```
Qy      2 GTCTGTCACAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCA-TT 60
          |||  |  ||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db    799 GTCCCTGGCAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTT 858

Qy     61 CTGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 100
          ||||||||||||||||||||||||||||||||||||||||
Db    859 CTGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 898


RESULT 9
AL521983/c
LOCUS       AL521983                820 bp    mRNA    linear   EST 13-FEB-2001
DEFINITION  AL521983 LTI_NFL004_NBC2 Homo sapiens cDNA clone CS0DB007YC01 3
            prime, mRNA sequence.
ACCESSION   AL521983
VERSION     AL521983.1  GI:12785476
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 820)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             Location/Qualifiers
     source          1. .820
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CS0DB007YC01"
                     /clone_lib="LTI_NFL004_NBC2"
                     /sex="male"
                     /tissue_type="neuroblastoma cells"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-stranded cDNA was digested with Not I and
                     cloned into the Not I and Eco RV sites of the pCMVSPORT 6
                     vector.  Library was normalized. Library was constructed
                     by Life Technologies.  Contact : Feng Liang Life
                     Technologies, a division of Invitrogen 9800 Medical Center
                     Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
                     8371 Email : fliang@lifetech.com URL :
                     http://fulllength.invitrogen.com"
BASE COUNT      121 a    222 c    260 g    184 t     33 others
ORIGIN

  Query Match            80.0%;  Score 80;  DB 9;  Length 820;
  Best Local Similarity  94.0%;  Pred. No. 1.1e-17;
  Matches   94;  Conservative    0;  Mismatches    5;  Indels    1;  Gaps    1;

Qy      2 GTCTGTCACAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGT-GGTGGAACCATT 60
          |||  |  |||||||||||||||||||||||||||||||||||||| |||||||| |||
Db    641 GTCCCTGGCAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGGTGGAACTATT 582

Qy     61 CTGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 100
          ||||||||||||||||||||||||||||||||||||||||
Db    581 CTGAGCGAGT"CTACATCCTAACGGCAGCCCACTGTCTC" 542


RESULT 10
AL521984
LOCUS       AL521984                898 bp    mRNA    linear   EST 13-FEB-2001
DEFINITION  AL521984 LTI_NFL004_NBC2 Homo sapiens cDNA clone CS0DB007YC01 5
            prime, mRNA sequence.
ACCESSION   AL521984
VERSION     AL521984.1  GI:12785477
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 898)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             Location/Qualifiers
     source          1. .898
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CS0DB007YC01"
                     /clone_lib="LTI_NFL004_NBC2"
                     /sex="male"
                     /tissue_type="neuroblastoma cells"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-stranded cDNA was digested with Not I and
                     cloned into the Not I and Eco RV sites of the pCMVSPORT 6
                     vector.  Library was normalized. Library was constructed
                     by Life Technologies.  Contact : Feng Liang Life
                     Technologies, a division of Invitrogen 9800 Medical Center
                     Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
                     8371 Email : fliang@lifetech.com URL :
                     http://fulllength.invitrogen.com"
BASE COUNT      234 a    253 c    264 g    147 t
ORIGIN

  Query Match            79.0%;  Score 79;  DB 9;  Length 898;
  Best Local Similarity  93.9%;  Pred. No. 2.7e-17;
  Matches   93;  Conservative    0;  Mismatches    5;  Indels    1;  Gaps    1;

Qy      2 GTCTGTCACAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 61
          |||  |  ||||||||||||||||||||||||||||||||||||||||||||||||| |||| |
Db    757 GTCCCTGGCAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGG-ACTATTC 815

Qy     62 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 100
          ||||||||||||||||||||||||||||||||||||||||
Db    816 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 854


RESULT 11
BG609107
LOCUS       BG609107                521 bp    mRNA    linear   EST 17-APR-2001
DEFINITION  322632 MARC 1PIG Sus scrofa cDNA 5', mRNA sequence.
ACCESSION   BG609107
VERSION     BG609107.1  GI:13659085
KEYWORDS    EST.
SOURCE      pig.
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 521)
  AUTHORS   Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,
            Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W.
            and Keele,J.W.
  TITLE     Design and use of two pooled tissue normalized cDNA libraries for
            EST discovery in swine
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@email.marc.usda.gov
            Single pass sequencing. Bases called and alt_trimmed with phred
            v0.980904.e. Vector identified by cross_match with the -minscore 18
            and -minmatch 12 options.
```

```
            PCR PRimers
            FORWARD: AGGAAACAGCTATGACCAT
            BACKWARD: GTTTTCCCAGTCACGACG
            Plate: 97  row: M  column: 13
            Seq primer: ATTTAGGTGACACTATAG.
FEATURES              Location/Qualifiers
     source           1. .521
                      /organism="Sus scrofa"
                      /db_xref="taxon:9823"
                      /clone_lib="MARC 1PIG"
                      /tissue_type="pooled"
                      /lab_host="DH10B"
                      /note="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI;
                      Library made from pooled tissue from day 11, 13, 15, 20,
                      and 30 embryos."
BASE COUNT      111 a    164 c    182 g     64 t
ORIGIN

  Query Match             70.8%;  Score 70.8;  DB 12;  Length 521;
  Best Local Similarity   86.7%;  Pred. No. 1.7e-14;
  Matches   78;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy      10 CAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCA TTCTCAGCGAG 69
           |||||||||||| |||| ||||||||||||||| |||||||| || ||||| |||||||||
Db      91 CAGGCCCTGCTCGTCAACGAGGAAAACGAGGGCTTCTGTGGCGGCACCATCCTGAGCGAG 150

Qy      70 TTCTACATCCTAACGGCAGCCCACTGTCTC 99
           ||| || | || ||||| |||||||| |||
Db     151 TTCCACGTGCTCACGGCGGCCCACTGCCTC 180


RESULT 12
W70576
LOCUS       W70576                   609 bp    mRNA    linear   EST 17-JUN-1996
DEFINITION  me42a04.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA
            clone IMAGE:390126 5' similar to gb:M57285 COAGULATION FACTOR X
            PRECURSOR (HUMAN);, mRNA sequence.
ACCESSION   W70576
VERSION     W70576.1  GI:1380821
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 609)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:241958
            Seq primer: ETPrimer
            High quality sequence stop: 342.
FEATURES              Location/Qualifiers
     source           1. .609
                      /organism="Mus musculus"
                      /strain="C57BL/6J"
                      /db_xref="taxon:10090"
                      /clone="IMAGE:390126"
                      /clone_lib="Soares mouse embryo NbME13.5 14.5"
                      /sex="unknown"
                      /tissue_type="embryo"
                      /dev_stage="13.5-14.5dpc total fetus"
                      /lab_host="DH10B"
                      /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                      polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                      was primed with a Not I - oligo(dT) primer [5'
                      TGTTACCAATCTGAAGTGGGAGCGGCCGCGGAAATTTTTTTTTTTTTTTTTTT
                      T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
                      14.5dpc embryos [total RNA provided by  Minoru Ko, Wayne
                      State Univ., from 2 ]; double-stranded cDNA was ligated to
                      Eco RI adaptors (Pharmacia), digested with Not I and
                      cloned into the Not I and Eco RI sites of the modified
                      pT7T3 vector. Library went through one round of
                      normalization, and was constructed by Bento Soares and
                      M.Fatima Bonaldo.  "
BASE COUNT      157 a    150 c    177 g    125 t
ORIGIN

  Query Match             64.4%;  Score 64.4;  DB 14;  Length 609;
  Best Local Similarity   78.6%;  Pred. No. 3.7e-12;
  Matches   77;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy       2 GTCTGTCACAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 61
           |||  |  ||||| |||||||| || || || || || || |||||||| || |||||
Db     142 GTCCCTGGCAGGCTCTGCTCATTAACGAAGACAATGAAGGGTTCTGTGGGGGCACCATCT 201

Qy      62 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTC 99
           |||  |||||||||||||| || || ||||||||||||
Db     202 TGAATGAGTTCTACATCCTCACTGCTGCCCACTGTCTC 239


RESULT 13
W70579
LOCUS       W70579                   617 bp    mRNA    linear   EST 17-JUN-1996
DEFINITION  me42b03.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA
            clone IMAGE:390125 5' similar to gb:M57285 COAGULATION FACTOR X
            PRECURSOR (HUMAN);, mRNA sequence.
ACCESSION   W70579
VERSION     W70579.1  GI:1380824
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 617)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:241957
            Seq primer: ETPrimer
            High quality sequence stop: 343.
FEATURES              Location/Qualifiers
     source           1. .617
                      /organism="Mus musculus"
                      /strain="C57BL/6J"
                      /db_xref="taxon:10090"
                      /clone="IMAGE:390125"
                      /clone_lib="Soares mouse embryo NbME13.5 14.5"
                      /sex="unknown"
                      /tissue_type="embryo"
                      /dev_stage="13.5-14.5dpc total fetus"
```

/lab_host="DH10B"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGGCCGCGGAAATTTTTTTTTTTTTTTTTTTTTT T 3'], on equal amounts of mRNA from 2 13.5dpc and 2 14.5dpc embryos [total RNA provided by Minoru Ko, Wayne State Univ., from 2]; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M.Fatima Bonaldo. "

BASE COUNT 159 a 155 c 178 g 125 t
ORIGIN

Query Match 64.4%; Score 64.4; DB 14; Length 617;
Best Local Similarity 78.6%; Pred. No. 3.7e-12;
Matches 77; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

```
Qy      2 GTCTGTCACAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 61
          |||  |  ||||| |||||||| || || || || || || |||||||| || |||||
Db    142 GTCCCTGGCAGGCTCTGCTCATTAACGAAGACAATGAAGGGTTCTGTGGGGGCACCATCT 201

Qy     62 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTC 99
          |||  |||||||||||||| || || ||||||||||||
Db    202 TGAATGAGTTCTACATCCTCACTGCTGCCCACTGTCTC 239
```

RESULT 14
BI144826
LOCUS BI144826 824 bp mRNA linear EST 05-JUL-2001
DEFINITION 602910055F1 NCI_CGAP_Li9 Mus musculus cDNA clone IMAGE:5051317 5', mRNA sequence.
ACCESSION BI144826
VERSION BI144826.1 GI:14604827
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 824)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11139 row: m column: 14
High quality sequence stop: 816.
FEATURES Location/Qualifiers
source 1. .824
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5051317"
/clone_lib="NCI_CGAP_Li9"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: liver; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.9 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 212 a 204 c 248 g 160 t
ORIGIN

Query Match 64.4%; Score 64.4; DB 13; Length 824;
Best Local Similarity 78.6%; Pred. No. 4.3e-12;
Matches 77; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

```
Qy      2 GTCTGTCACAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 61
          |||  |  ||||| |||||||| || || || || || || |||||||| || |||||
Db    145 GTCCCTGGCAGGCTCTGCTCATTAACGAAGACAATGAAGGGTTCTGTGGGGGCACCATCT 204

Qy     62 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTC 99
          |||  |||||||||||||| || || ||||||||||||
Db    205 TGAATGAGTTCTACATCCTCACTGCTGCCCACTGTCTC 242
```

RESULT 15
AA237419
LOCUS AA237419 582 bp mRNA linear EST 03-MAR-1997
DEFINITION mx19d10.r1 Soares mouse NML Mus musculus cDNA clone IMAGE:680659 5' similar to gb:M57285 COAGULATION FACTOR X PRECURSOR (HUMAN);, mRNA sequence.
ACCESSION AA237419
VERSION AA237419.1 GI:1861498
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 582)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
TITLE The WashU-HHMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:42063
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 486.
FEATURES Location/Qualifiers
source 1. .582
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:680659"
/clone_lib="Soares mouse NML"
/tissue_type="Liver"
/lab_host="DH10B"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGGCCGCGAATCTTTTTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library constructed and normalized by Bento Soares and M.Fatima Bonaldo."
BASE COUNT 147 a 148 c 173 g 114 t
ORIGIN

Query Match 58.8%; Score 58.8; DB 9; Length 582;
Best Local Similarity 80.2%; Pred. No. 3.6e-10;
Matches 69; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

```
Qy     14 CCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTCTGAGCGAGTTCT 73
          | ||||||||| || || || || || || |||||||| || |||||  |||  |||||||
Db      1 CTCTGCTCATTAACGAAGACAATGAAGGGTTCTGTGGGGGCACCATCTTGAATGAGTTCT 60

Qy     74 ACATCCTAACGGCAGCCCACTGTCTC 99
```

```
                 ||||||| || || ||||||||||||
Db       61 ACATCCTCACTGCTGCCCACTGTCTC 86
```

Search completed: January 15, 2003, 21:18:51
Job time : 1272 secs

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: January 15, 2003, 17:56:05 ; Search time 1334 Seconds
(without alignments)
2181.620 Million cell updates/sec

Title: L00396_COPY_1_100
Perfect score: 100
Sequence: 1 CGTCTGTCACAGGCCCTGCT.....AACGGCAGCCCACTGTCTCT 100

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	100	100.0	142	9	HUMFX7	L00396 Human facto
2	100	100.0	132933	9	AL137002	AL137002 Human DNA
3	98.4	98.4	29488	9	AF503510	AF503510 Homo sapi
4	92.6	92.6	725	6	AR121387	AR121387 Sequence
5	92.6	92.6	725	6	AR124119	AR124119 Sequence
6	92.6	92.6	725	6	AX022601	AX022601 Sequence
7	92.6	92.6	725	6	E36142	E36142 Chimeric se
8	92.6	92.6	1126	6	AR095306	AR095306 Sequence
9	92.6	92.6	1126	6	AR103990	AR103990 Sequence
10	92.6	92.6	1126	9	HUMFX	K01886 Human facto
11	92.6	92.6	1404	6	A93124	A93124 Sequence 15
12	92.6	92.6	1414	9	HUMCFX	M22613 Human blood
13	91	91.0	1443	9	HUMFXM	K03194 Human facto
14	91	91.0	1467	6	A86859	A86859 Sequence 43
15	91	91.0	1467	6	A86886	A86886 Sequence 26
16	91	91.0	1467	6	AX082959	AX082959 Sequence
17	91	91.0	1500	6	AR024194	AR024194 Sequence
18	91	91.0	1507	9	HUMFACX	M57285 Human coagu
19	66.8	66.8	20130	10	AF211347	AF211347 Mus muscu
20	66.8	66.8	299880	2	AC127308	AC127308 Mus muscu
21	64.4	64.4	1486	10	AF087644	AF087644 Mus muscu
22	64.4	64.4	1491	10	MMCOAGULX	AJ222677 Mus muscu
23	63.4	63.4	1537	4	BTCFXI	X00673 Bovine mRNA
24	63.4	63.4	1554	6	A73583	A73583 Sequence 1
25	63.4	63.4	1554	6	AR001423	AR001423 Sequence
26	62.8	62.8	1925	10	BC003877	BC003877 Mus muscu
27	61.8	61.8	1500	4	AF003200	AF003200 Oryctolag
28	55.8	55.8	20947	5	AF519546	AF519546 Danio rer
29	48.8	48.8	1497	10	RNFXRAT	X79807 R.norvegicu
30	42.8	42.8	2246	5	CHKVAP	D00844 Gallus gall
31	42.6	42.6	1599	4	AF275654	AF275654 Ornithorh
c 32	37.4	37.4	164553	2	AC129680	AC129680 Rattus no
33	37.4	37.4	192453	2	AC129457	AC129457 Rattus no
34	37.2	37.2	855	10	GPIFIXA	M26237 Guinea pig
35	36	36.0	846	10	RATFIXA	M26247 Rat factor
36	36	36.0	852	10	MUSFIXA	M26236 Mouse facto
37	36	36.0	1047	6	AX250084	AX250084 Sequence
38	36	36.0	2697	10	MUSFIX	M23109 Mouse coagu
39	36	36.0	219448	2	AL671984	AL671984 Mus muscu
40	35.2	35.2	111516	5	AL672083	AL672083 Zebrafish
41	34.4	34.4	1360	3	AY118979	AY118979 Drosophil
c 42	34.4	34.4	95613	2	AC020389	AC020389 Drosophil
c 43	34.4	34.4	155840	2	AC006495	AC006495 Drosophil
c 44	34.4	34.4	199044	3	AC009394	AC009394 Drosophil
c 45	34.4	34.4	224896	3	AE003721	AE003721 Drosophil

ALIGNMENTS

RESULT 1
HUMFX7
LOCUS HUMFX7 142 bp DNA linear PRI 09-NOV-1994
DEFINITION Human factor X (blood coagulation factor) gene, exon 7.
ACCESSION L00396 M14327
VERSION L00396.1 GI:182828
KEYWORDS Stuart factor; blood coagulation factor; factor X; glycoprotein; serine protease.
SEGMENT 7 of 8
SOURCE Homo sapiens (tissue library: of Lawn et al., and Yoshitake et al.) DNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 142)

```
   AUTHORS   Leytus,S.P., Foster,D.C., Kurachi,K. and Davie,E.W.
   TITLE     Gene for human factor X: a blood coagulation factor whose gene
             organization is essentially identical with that of factor IX and
             protein C
   JOURNAL   Biochemistry 25 (18), 5098-5102 (1986)
   MEDLINE   87026600
    PUBMED   3768336
COMMENT      matp    +    13  +   130     factor Xa heavy chain.
FEATURES             Location/Qualifiers
     source          1. .142
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /map="13q34"
                     /tissue_lib="of Lawn et al., and Yoshitake et al."
     gene            join(L00390.1:1. .95,L00391.1:1. .185,L00392.1:1. .49,
                     L00393.1:1. .138,L00394.1:1. .156,L00395.1:1. .269,1. .130)
                     /gene="F10"
     intron          <1. .12
                     /gene="F10"
                     /note="FX intron F"
     exon            13. .130
                     /gene="F10"
                     /note="G00-119-890"
                     /number=7
     intron          131. .>142
                     /note="FX intron G"
BASE COUNT       38 a     35 c     38 g     31 t
ORIGIN      About 3.4 kb after segment 6; chromosome 13q34.

  Query Match             100.0%;  Score 100;  DB 9;  Length 142;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-24;
  Matches  100;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CGTCTGTCACAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATT 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CGTCTGTCACAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATT 60

Qy      61 CTGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 100
           |||||||||||||||||||||||||||||||||||||||
Db      61 CTGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 100


RESULT 2
AL137002
LOCUS       AL137002               132933 bp    DNA     linear   PRI 04-MAY-2001
DEFINITION  Human DNA sequence from clone RP11-98F14 on chromosome
            13q22.1-31.1, complete sequence.
ACCESSION   AL137002
 VERSION    AL137002.19  GI:13990998
  KEYWORDS  HTG.
  SOURCE    human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 132933)
  AUTHORS   Bates,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-MAY-2001) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT     On May 8, 2001 this sequence version replaced gi:13274222.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest. The following
            abbreviations are used to associate primary accession numbers given
            in the feature table with their source databases: Em:, EMBL; Sw:,
            SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP
            database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 13, constructed by the Sanger Centre Chromosome 13
            Mapping Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr13
            RP11-98F14 is from the library RPCI-11.1 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6
            IMPORTANT: This sequence is not the entire insert of clone
            RP11-98F14 It may be shorter because we sequence overlapping
            sections only once, except for a 100 base overlap.
            The true left end of clone RP11-98F14 is at 1 in this sequence. The
            true left end of clone RP11-391H12 is at 132834 in this sequence.
            The true right end of clone RP11-265C7 is at 123923 in this
            sequence.
FEATURES             Location/Qualifiers
     source          1. .132933
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="13"
                     /map="q22.1-31.1"
                     /clone="RP11-98F14"
                     /clone_lib="RPCI-11.1"
     repeat_region   41. .277
                     /note="AluJo repeat: matches 1. .229 of consensus"
     repeat_region   377. .472
                     /note="3 copies 32 mer 93% conserved"
     repeat_region   901. .1207
                     /note="AluSq repeat: matches 1. .307 of consensus"
     misc_feature    1483. .1912
                     /note="CpG island"
                     /evidence=not_experimental
     repeat_region   1999. .2164
                     /note="MER58A repeat: matches 19. .167 of consensus"
     repeat_region   2703. .2861
                     /note="L1M3e repeat: matches -118. .34 of consensus"
     repeat_region   4371. .4978
                     /note="38 copies 16 mer 78% conserved"
     repeat_region   4374. .4981
                     /note="19 copies 32 mer 87% conserved"
     repeat_region   6373. .6448
                     /note="2 copies 38 mer 93% conserved"
     repeat_region   6411. .6563
                     /note="3 copies 51 mer 93% conserved"
     repeat_region   6565. .6669
                     /note="3 copies 35 mer 85% conserved"
     repeat_region   6612. .6687
                     /note="2 copies 38 mer 98% conserved"
     repeat_region   6672. .6773
                     /note="2 copies 51 mer 93% conserved"
     repeat_region   8852. .8923
                     /note="4 copies 18 mer 81% conserved"
     repeat_region   8854. .8988
                     /note="15 copies 9 mer cccgaaccc 63% conserved"
     repeat_region   8860. .8943
                     /note="3 copies 28 mer 83% conserved"
     repeat_region   8920. .9059
                     /note="4 copies 35 mer 75% conserved"
     repeat_region   8927. .9056
                     /note="5 copies 26 mer 69% conserved"
     repeat_region   8935. .9002
                     /note="4 copies 17 mer 95% conserved"
     repeat_region   8984. .9037
                     /note="3 copies 18 mer 90% conserved"
     repeat_region   9427. .9530
                     /note="4 copies 26 mer 92% conserved"
     repeat_region   10204. .10602
```

```
                     /note="7 copies 57 mer 87% conserved"
     repeat_region   11050. .11091
                     /note="2 copies 21 mer 100% conserved"
     repeat_region   11469. .11823
                     /note="L1MC1 repeat: matches 5862. .6231 of consensus"
     repeat_region   13157. .13200
                     /note="LTR5 repeat: matches 690. .733 of consensus"
     repeat_region   13806. .14155
                     /note="5 copies 70 mer 82% conserved"
     misc_feature    15491. .15973
                     /note="CpG island"
                     /evidence=not_experimental
     repeat_region   16278. .16371
                     /note="2 copies 47 mer 87% conserved"
     repeat_region   16956. .17159
                     /note="4 copies 51 mer 95% conserved"
     repeat_region   19411. .19587
                     /note="3 copies 59 mer 79% conserved"
     repeat_region   20202. .20342
                     /note="3 copies 47 mer 96% conserved"
     repeat_region   20473. .20564
                     /note="L2 repeat: matches 2602. .2693 of consensus"
     misc_feature    20816. .21894
                     /note="CpG island"
                     /evidence=not_experimental
     repeat_region   24570. .24680
                     /note="L1M4 repeat: matches 3584. .3686 of consensus"
     repeat_region   24681. .24999
                     /note="AluJo repeat: matches 1. .311 of consensus"
     repeat_region   25000. .25306
                     /note="L1M4 repeat: matches 3686. .4029 of consensus"
     repeat_region   25311. .25620
                     /note="AluY repeat: matches 1. .311 of consensus"
     misc_feature    25345. .25989
                     /note="CpG island"
                     /evidence=not_experimental
     repeat_region   25755. .26053
                     /note="AluSq repeat: matches 1. .304 of consensus"
     repeat_region   27585. .27688
                     /note="4 copies 26 mer 74% conserved"
     repeat_region   27733. .27981
                     /note="AluJo repeat: matches 9. .282 of consensus"
     repeat_region   29431. .29498
                     /note="MER81 repeat: matches 4. .72 of consensus"
     repeat_region   29553. .29870
                     /note="AluSx repeat: matches 1. .309 of consensus"
     repeat_region   30907. .31396
                     /note="7 copies 70 mer 64% conserved"
     repeat_region   30908. .31432
                     /note="15 copies 35 mer 61% conserved"
     repeat_region   30911. .31412
                     /note="251 copies 2 mer gg 61% conserved"
     repeat_region   30938. .31413
                     /note="17 copies 28 mer 61% conserved"
     repeat_region   30973. .31420
                     /note="14 copies 32 mer 61% conserved"
     repeat_region   30987. .31385
                     /note="7 copies 57 mer 64% conserved"
     misc_feature    31000. .31080
                     /note="Sequence from uni-directional dGTP big dye
                     terminator reads only."
     repeat_region   31043. .31413
                     /note="7 copies 53 mer 63% conserved"
     repeat_region   31078. .31415
                     /note="13 copies 26 mer 62% conserved"
     repeat_region   31136. .31417
                     /note="6 copies 47 mer 66% conserved"
     repeat_region   31152. .31421
                     /note="9 copies 30 mer 61% conserved"
     misc_feature    31153. .31267
                     /note="Single clone region. Sequence from reads from a
                     short insert library derived from a single pUC clone.
                     Restriction digest data confirm the assembly."
     repeat_region   31175. .31426
                     /note="12 copies 21 mer 61% conserved"
     repeat_region   31268. .31411
                     /note="9 copies 16 mer 64% conserved"
     misc_feature    31268
                     /note="Tandem repeat. Forced join. Gap size estimated to
                     be approximately 350bp by restriction digest data."
     repeat_region   31882. .32952
                     /note="21 copies 51 mer 89% conserved"
     repeat_region   31884. .32937
                     /note="62 copies 17 mer 89% conserved"
     misc_feature    31896. .35863
                     /note="CpG island"
                     /evidence=not_experimental
     repeat_region   31982. .33941
                     /note="28 copies 70 mer 54% conserved"
     repeat_region   32243. .32434
                     /note="6 copies 32 mer 64% conserved"
     repeat_region   32271. .32535
                     /note="5 copies 53 mer 64% conserved"
     repeat_region   32883. .32952
                     /note="2 copies 35 mer 92% conserved"
     repeat_region   33336. .33475
                     /note="4 copies 35 mer 70% conserved"
     repeat_region   33367. .33893

  Query Match             100.0%;  Score 100;  DB 9;  Length 132933;
  Best Local Similarity   100.0%;  Pred. No. 3.5e-24;
  Matches  100;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CGTCTGTCACAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATT 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   72010 CGTCTGTCACAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATT 72069

Qy      61 CTGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 100
           |||||||||||||||||||||||||||||||||||||||
Db   72070 CTGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 72109


RESULT 3
AF503510
LOCUS       AF503510               29488 bp    DNA     linear   PRI 22-MAY-2002
DEFINITION  Homo sapiens coagulation factor X (F10) gene, complete cds.
ACCESSION   AF503510
VERSION     AF503510.1  GI:20336662
KEYWORDS    .
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 29488)
  AUTHORS   Rieder,M.J., Armel,T.Z., Carrington,D.P., Chung,M.-W., Lee,K.L.,
            Ozuna,M., Poel,C.L., Toth,E.J., Yi,Q. and Nickerson,D.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-APR-2002) Genome Sciences, University of Washington,
            1705 NE Pacific, Seattle, WA 98195, USA
COMMENT     To cite this work please use:  SeattleSNPs. NHLBI HL66682 Program
            for Genomic Applications, UW-FHCRC, Seattle, WA (URL:
            http://pga.gs.washington.edu).
FEATURES             Location/Qualifiers
     source          1. .29488
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     variation       152
                     /frequency="0.04"
                     /replace="c"
     variation       167
                     /frequency="0.01"
                     /replace="t"
     variation       370
                     /frequency="0.01"
                     /replace="a"
     variation       577
```

```
                     /frequency="0.33"
                     /replace="t"
     variation       671
                     /frequency="0.01"
                     /replace="t"
     variation       713
                     /frequency="0.12"
                     /replace="c"
     variation       851
                     /frequency="0.01"
                     /replace="g"
     variation       866
                     /frequency="0.02"
                     /replace="a"
     variation       880
                     /frequency="0.05"
                     /replace="c"
     variation       1116
                     /frequency="0.79"
                     /replace=""
     variation       1169
                     /frequency="0.02"
                     /replace="a"
     repeat_region   1189. .1395
                     /rpt_family="MIR"
                     /rpt_type=dispersed
     variation       1319. .1324
                     /frequency="0.57"
                     /replace=""
     variation       1449
                     /frequency="0.48"
                     /replace="c"
     variation       1451
                     /frequency="0.33"
                     /replace="a"
     variation       1612
                     /frequency="0.01"
                     /replace="a"
     variation       1632
                     /frequency="0.01"
                     /replace="t"
     gene            <1647. .28346
                     /gene="F10"
     mRNA            join(<1647. .1741,8269. .8429,17274. .17298,18174. .18287,
                     19736. .19867,22668. .22912,26198. .26315,27735. .28346)
                     /gene="F10"
                     /product="coagulation factor X"
     CDS             join(1672. .1741,8269. .8429,17274. .17298,18174. .18287,
                     19736. .19867,22668. .22912,26198. .26315,27735. .28336)
                     /gene="F10"
                     /codon_start=1
                     /product="coagulation factor X"
                     /protein_id="AAM19347.1"
                     /db_xref="GI:20336663"
                     /translation="MGRPLHLVLLSASLAGLLLLGESLFIRREQANNILARVTRANSF
                     LEEMKKGHLERECMEETCSYEEAREVFEDSDKTNEFWNKYKDGDQCETSPCQNQGKCK
                     DGLGEYTCTCLEGFEGKNCELFTRKLCSLDNGDCDQFCHEEQNSVVCSCARGYTLADN
                     GKACIPTGPYPCGKQTLERRKRSVAQATSSSGEAPDSITWKPYDAADLDPTENPFDLL
                     DFNQTQPERGDNNLTRIVGGQECKDGECPWQALLINEENEGFCGGTILSEFYILTAAH
                     CLYQAKRFKVRVGDRNTEQEEGGEAVHEVEVVIKHNRFTKETYDFDIAVLRLKTPITF
                     RMNVAPACLPERDWAESTLMTQKTGIVSGFGRTHEKGRQSTRLKMLEVPYVDRNSCKL
                     SSSFIITQNMFCAGYDTKQEDACQGDSGGPHVTRFKDTYFVTGIVSWGEGCARKGKYG
                     IYTKVTAFLKWIDRSMKTRGLPKAKSHAPEVITSSPLK"
     variation       1731
                     /gene="F10"
                     /frequency="0.01"
                     /replace="t"
     variation       2011
                     /gene="F10"
                     /frequency="0.22"
                     /replace="g"
     variation       2456
                     /gene="F10"
                     /frequency="0.04"
                     /replace="a"
     variation       2527
                     /gene="F10"
                     /frequency="0.12"
                     /replace="c"
     variation       2757
                     /gene="F10"
                     /frequency="0.04"
                     /replace="g"
     misc_feature    2860. .3410
                     /gene="F10"
                     /note="Region not scanned for variation"
     repeat_region   3544. .3975
                     /rpt_family="MaLR"
                     /rpt_type=dispersed
     variation       3715
                     /gene="F10"
                     /frequency="0.09"
                     /replace="t"
     variation       3740
                     /gene="F10"
                     /frequency="0.39"
                     /replace="g"
     variation       3977
                     /gene="F10"
                     /frequency="0.01"
                     /replace="c"
     variation       4100
                     /gene="F10"
                     /frequency="0.01"
                     /replace="t"
     variation       4544
                     /gene="F10"
                     /frequency="0.19"
                     /replace="c"
     repeat_region   4567. .4967
                     /rpt_family="L2"
                     /rpt_type=dispersed
     variation       5128
                     /gene="F10"
                     /frequency="0.01"
                     /replace="g"
     variation       5300
                     /gene="F10"
                     /frequency="0.02"
                     /replace="c"
     variation       5413
                     /gene="F10"
                     /frequency="0.13"
                     /replace="c"
     variation       5521
                     /gene="F10"
                     /frequency="0.11"
                     /replace="a"
     misc_feature    5540. .5950
                     /gene="F10"
                     /note="Region not scanned for variation"
     variation       5997
                     /gene="F10"
                     /frequency="0.01"
                     /replace="c"
     variation       6098
                     /gene="F10"
                     /frequency="0.03"
                     /replace="a"
     variation       6110
                     /gene="F10"
                     /frequency="0.01"
                     /replace="c"
     variation       6445
                     /gene="F10"
                     /frequency="0.06"
```

```
                     /replace="c"
     variation       6638. .6641
                     /gene="F10"
                     /frequency="0.99"
                     /replace=""
     repeat_region   6943. .7395
                     /rpt_family="ERV1"
                     /rpt_type=dispersed
     repeat_region   7396. .7727
                     /rpt_family="Alu"
                     /rpt_type=dispersed
     variation       7437
                     /gene="F10"
                     /frequency="0.01"
                     /replace="a"
     repeat_region   7728. .8145
                     /rpt_family="ERV1"
                     /rpt_type=dispersed
     variation       7742
                     /gene="F10"
                     /frequency="0.01"
                     /replace="t"

  Query Match             98.4%;  Score 98.4;  DB 9;  Length 29488;
  Best Local Similarity   99.0%;  Pred. No. 1.2e-23;
  Matches   99;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CGTCTGTCACAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATT 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db   26186 CGTCTGTCACAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACTATT 26245

Qy      61 CTGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 100
           |||||||||||||||||||||||||||||||||||||||||
Db   26246 CTGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 26285


RESULT 4
AR121387
LOCUS       AR121387                725 bp    DNA     linear   PAT 16-MAY-2001
DEFINITION  Sequence 12 from patent US 6159722.
ACCESSION   AR121387
VERSION     AR121387.1  GI:14104963
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 725)
  AUTHORS   Bode,W., Engh,R., Hopfner,K.-P., Huber,R. and Kopetzki,E.
  TITLE     Chimeric serine proteases
  JOURNAL   Patent: US 6159722-A 12 12-DEC-2000;
FEATURES             Location/Qualifiers
     source          1. .725
                     /organism="unknown"
BASE COUNT      172 a    198 c    216 g    139 t
ORIGIN

  Query Match             92.6%;  Score 92.6;  DB 6;  Length 725;
  Best Local Similarity   96.0%;  Pred. No. 1e-21;
  Matches   95;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 GTCTGTCACAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 61
           |||  |  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      65 GTCCCTGGCAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 124

Qy      62 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 100
           |||||||||||||||||||||||||||||||||||||||
Db     125 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 163


RESULT 5
AR124119
LOCUS       AR124119                725 bp    DNA     linear   PAT 16-MAY-2001
DEFINITION  Sequence 12 from patent US 6171842.
ACCESSION   AR124119
VERSION     AR124119.1  GI:1'109480
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 725)
  AUTHORS   Bode,W., Engh,R., Hopfner,K.-P., Huber,R. and Kopetzki,E.
  TITLE     Chimeric serine proteases
  JOURNAL   Patent: US 6171842-A 12 09-JAN-2001;
FEATURES             Location/Qualifiers
     source          1. .725
                     /organism="unknown"
BASE COUNT      172 a    198 c    216 g    139 t
ORIGIN

  Query Match             92.6%;  Score 92.6;  DB 6;  Length 725;
  Best Local Similarity   96.0%;  Pred. No. 1e-21;
  Matches   95;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 GTCTGTCACAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 61
           |||  |  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      65 GTCCCTGGCAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 124

Qy      62 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 100
           |||||||||||||||||||||||||||||||||||||||
Db     125 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 163


RESULT 6
AX022601
LOCUS       AX022601                725 bp    DNA     linear   PAT 07-SEP-2000
DEFINITION  Sequence 12 from Patent EP0927764.
ACCESSION   AX022601
VERSION     AX022601.1  GI:10046158
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 725)
  AUTHORS   Bode,W., Engh,R., Huber,R., Kopetzki,E. and Hopfner,K.P.
  TITLE     Chimeric serine proteases
  JOURNAL   Patent: EP 0927764-A 12 07-JUL-1999;
            ROCHE DIAGNOSTICS GMBH (DE)
FEATURES             Location/Qualifiers
     source          1. .725
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT      172 a    198 c    216 g    139 t
ORIGIN

  Query Match             92.6%;  Score 92.6;  DB 6;  Length 725;
  Best Local Similarity   96.0%;  Pred. No. 1e-21;
  Matches   95;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 GTCTGTCACAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 61
           |||  |  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      65 GTCCCTGGCAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 124

Qy      62 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 100
           |||||||||||||||||||||||||||||||||||||||
Db     125 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 163


RESULT 7
E36142
LOCUS       E36142                  725 bp    DNA     linear   PAT 18-JUN-2001
DEFINITION  Chimeric serine protease.
ACCESSION   E36142
VERSION     E36142.1  GI:13022520
KEYWORDS    JP 1999235173-A/12.
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
```

```
            artificial sequences.
REFERENCE   1  (bases 1 to 725)
  AUTHORS   Wolfran,B., Richard,E., Karl,P.H., RoLert,H. and Erhard,K.
  TITLE     Chimeric serine protease
  JOURNAL   Patent: JP 1999235173-A 12 31-AUG-1999;
            ROCHE DIAGNOSTICS GMBH
COMMENT     OS   Artificial Sequence
            PN   JP 1999235173-A/12
            PD   31-AUG-1999
            PF   03-DEC-1998 JP 1998343777
            PR   03-DEC-1997 EP   97121232.9
            PI   WOLFRAN BODO,RICHARD ENGU,KARL PATER HOFUNER,ROBERT HYUBA, PI
             ERHARD KOPEKI
            PC   C12N9/50,C12N1/21,C12N15/09//(C12N9/50,C12R1:19),(C12N1/21, PC
              C12R1:19),
            PC   (C12N15/09,C12R1:91),C12N15/00,(C12N15/00,C12R1:91) CC
            FH   Key             Location/Qualifiers
            FT   source          1. .725
            FT                   /organism='Artificial Sequence'.
FEATURES              Location/Qualifiers
     source           1. .725
                      /organism="synthetic construct"
                      /db_xref="taxon:32630"
BASE COUNT      172 a    198 c    216 g    139 t
ORIGIN

  Query Match             92.6%;  Score 92.6;  DB 6;  Length 725;
  Best Local Similarity   96.0%;  Pred. No. 1e-21;
  Matches   95;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 GTCTGTCACAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 61
           |||  |  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      65 GTCCCTGGCAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 124

Qy      62 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 100
           ||||||||||||||||||||||||||||||||||||||||
Db     125 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 163


RESULT 8
AR095306
LOCUS       AR095306                1126 bp    DNA     linear   PAT 08-SEP-2000
DEFINITION  Sequence 27 from patent US 6004555.
ACCESSION   AR095306
VERSION     AR095306.1  GI:10023064
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1126)
  AUTHORS   Thorpe,P.E. and Edgington,T.S.
  TITLE     Methods for the specific coagulation of vasculature
  JOURNAL   Patent: US 6004555-A 27 21-DEC-1999;
FEATURES              Location/Qualifiers
     source           1. .1126
                      /organism="unknown"
BASE COUNT      269 a    341 c    342 g    174 t
ORIGIN

  Query Match             92.6%;  Score 92.6;  DB 6;  Length 1126;
  Best Local Similarity   96.0%;  Pred. No. 1e-21;
  Matches   95;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 GTCTGTCACAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 61
           |||  |  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     395 GTCCCTGGCAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 454

Qy      62 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 100
           ||||||||||||||||||||||||||||||||||||||||
Db     455 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 493


RESULT 9
AR103990
LOCUS       AR103990                1126 bp    DNA     linear   PAT 14-FEB-2001
DEFINITION  Sequence 27 from patent US 6093399.
ACCESSION   AR103990
VERSION     AR103990.1  GI:12816698
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1126)
  AUTHORS   Thorpe,P.E. and Edgington,T.S.
  TITLE     Methods and compositions for the specific coagulation of
            vasculature
  JOURNAL   Patent: US 6093399-A 27 25-JUL-2000;
FEATURES              Location/Qualifiers
     source           1. .1126
                      /organism="unknown"
BASE COUNT      269 a    341 c    342 g    174 t
ORIGIN

  Query Match             92.6%;  Score 92.6;  DB 6;  Length 1126;
  Best Local Similarity   96.0%;  Pred. No. 1e-21;
  Matches   95;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 GTCTGTCACAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 61
           |||  |  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     395 GTCCCTGGCAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 454

Qy      62 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 100
           ||||||||||||||||||||||||||||||||||||||||
Db     455 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 493


RESULT 10
HUMFX
LOCUS       HUMFX                   1126 bp    mRNA    linear   PRI 08-NOV-1994
DEFINITION  Human factor X mRNA.
ACCESSION   K01886
VERSION     K01886.1  GI:182820
KEYWORDS    Stuart factor; factor X; serine protease.
SOURCE      Human liver, cDNA to mRNA, clone lambda-X-1137.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1126)
  AUTHORS   Leytus,S.P., Chung,D.W., Kisiel,W., Kurachi,K. and Davie,E.W.
  TITLE     Characterization of a cDNA coding for human factor X
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 81 (12), 3699-3702 (1984)
  MEDLINE   84222026
   PUBMED   6587384
COMMENT     In processing, factor X (Stuart factor) is converted to Xa by
            cleavage of a glycopeptide from the amino-terminal end of the heavy
            chain. It then acts as a serine protease in converting prothrombin
            to thrombin.
FEATURES              Location/Qualifiers
     source           1. .1126
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /map="13q34"
     gene             1. .1126
                      /gene="F10"
     mRNA             <1. .1126
                      /gene="F10"
                      /product="factor X mRNA"
     CDS              <1. .1116
                      /gene="F10"
                      /note="factor X precursor peptide"
                      /codon_start=1
                      /protein_id="AAA52486.1"
                      /db_xref="GI:182821"
                      /db_xref="GDB:G00-119-890"
                      /translation="GFEGKNCELFTRKLCSLDNGDCDQFCHEEQNSVVCSCARGYTLA
                      DNGKACIPTGPYPCGKQTLERRKRSVAQATSSSGEAPDSITWKPYDAADLDPTENPFD
```

```
                    LLDFNQTQPERGDNNLTRIVGGQECKDGECPWQALLINEENEGFCGGTILSEFYILTA
                    AHCLYQAKRFEGDRNTEQEEGGEAVHEVEVVIKHNRFTKETYDFDIAVLRLKTPITFR
                    MNVAPACLPERDWAESTLMTQKTGIVSGFGRTHEKGRQSTRLKMLEVPYVDRNSCKLS
                    SSFIITQNMFCAGYDTKQEDACQGDSGGPHVTRFKDTYFVTGIVSWGEGCARKGKYGI
                    YTKVTAFLKWIDRSMKTRGLPKAKSHAPEVITSSPLK"
     mat_peptide    <1. .195
                    /gene="F10"
                    /product="factor X light chain"
     mat_peptide    205. .1113
                    /gene="F10"
                    /product="factor X heavy chain"
     mat_peptide    361. .1113
                    /gene="F10"
                    /product="factor Xa heavy chain"
BASE COUNT      269 a    341 c    342 g    174 t
ORIGIN      5 bp upstream of TaqI site.

  Query Match             92.6%;  Score 92.6;  DB 9;  Length 1126;
  Best Local Similarity   96.0%;  Pred. No. 1e-21;
  Matches   95;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 GTCTGTCACAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 61
           |||  |  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     395 GTCCCTGGCAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 454

Qy      62 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 100
           ||||||||||||||||||||||||||||||||||||||||
Db     455 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 493


RESULT 11
A93124
LOCUS       A93124                  1404 bp    DNA     linear   PAT 22-JAN-2000
DEFINITION  Sequence 15 from Patent WO9747737.
ACCESSION   A93124
VERSION     A93124.1  GI:6741514
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 1404)
  AUTHORS   Kopetzki,E. and Hopfner,K.
  TITLE     RECOMBINANT BLOOD-COAGULATION PROTEASES
  JOURNAL   Patent: WO 9747737-A 15 18-DEC-1997;
            KOPETZKI ERHARD (DE); BOEHRINGER MANNHEIM GMBH (DE)
FEATURES             Location/Qualifiers
     source          1. .1404
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT      356 a    402 c    425 g    221 t
ORIGIN

  Query Match             92.6%;  Score 92.6;  DB 6;  Length 1404;
  Best Local Similarity   96.0%;  Pred. No. 1.1e-21;
  Matches   95;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 GTCTGTCACAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 61
           |||  |  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     683 GTCCCTGGCAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 742

Qy      62 TGAGCGAC TCTACATCCTAACGGCAGCCCACTGTCT T 100
           ||||||||||||||||||||||||||||||||||||||||
Db     743 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 781


RESULT 12
HUMCFX
LOCUS       HUMCFX                  1414 bp    mRNA    linear   PRI 01-NOV-1994
DEFINITION  Human blood-coagulation factor X mRNA, complete cds.
ACCESSION   M22613
VERSION     M22613.1  GI:180335
KEYWORDS    coagulation factor X.
SOURCE      Human liver, cDNA to mRNA, clone pKT218.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1414)
  AUTHORS   Kaul,R.K., Hildebrand,B., Roberts,S. and Jagadeeswaran,P.
  TITLE     Isolation and characterization of human blood-coagulation factor X
            cDNA
  JOURNAL   Gene 41 (2-3), 311-314 (1986)
  MEDLINE   86221713
   PUBMED   3011603
FEATURES             Location/Qualifiers
     source          1. .1414
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /map="13q34"
     gene            1. .1414
                     /gene="F10"
     mRNA            <1. .1414
                     /gene="F10"
                     /product="coagulation factor X mRNA"
     CDS             <1. .1404
                     /gene="F10"
                     /note="coagulation factor X precursor"
                     /codon_start=1
                     /protein_id="AAA51984.1"
                     /db_xref="GI:180336"
                     /db_xref="GDB:G00-119-890"
                     /translation="LLGESLFIRREQANNILARVTRANSFLEEMKKGHLERECMEETC
                     SYEEAREVFEDSDKTNEFWNKYKDGDQCETSPCQNQGKCKDGLGEYTCTCLEGFEGKN
                     CELFTRKLCSLDNGDCDQFCHEEQNSVVCSCARGYTLADNGKACIPTGPYPCGKQTLE
                     RRKRSVAQATSSSGEAPDSITWKPYDAADLDPTENPFDLLDFNQTQPERGDNNLTRIV
                     GGQECKDGECPWQALLINEENEGFCGGTILSEFYILTAAHCLYQAKRFEGDRNTEQEE
                     GGEAVHEVEVVIKHNRFTKETYDFDIAVLRLKTPITFRMNVAPACLPERDWAESTLMT
                     QKTGIVSGFGRTHEKGRQSTRLKMLEVPYVDRNSCKLSSSFIITQNMFCAGYDTKQED
                     ACQGDSGGPHVTRFKDTYFVTGIVSWGEGCARKGKYGIYTKVTAFLKWIDRSMKTRGL
                     PKAKSHAPEVITSSPLK"
     sig_peptice     <1. .66
                     /gene="F10"
                     /note="coagulation factor X signal peptide"
     mat_peptide     67. .483
                     /gene="F10"
                     /product="coagulation factor X light chain"
     mat_peptide     493. .1401
                     /gene="F10"
                     /product="coagulation factor X heavy chain"
     mat_peptide     493. .648
                     /gene="F10"
                     /product="coagulation factor X activation peptide"
BASE COUNT      358 a    407 c    426 g    223 t
ORIGIN

  Query Match             92.6%;  Score 92.6;  DB 9;  Length 1414;
  Best Local Similarity   96.0%;  Pred. No. 1.1e-21;
  Matches   95;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 GTCTGTCACAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 61
           |||  |  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     683 GTCCCTGGCAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 742

Qy      62 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 100
           ||||||||||||||||||||||||||||||||||||||||
Db     743 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 781


RESULT 13
HUMFXM
LOCUS       HUMFXM                  1443 bp    mRNA    linear   PRI 08-NOV-1994
DEFINITION  Human factor X mRNA, partial signal pept and complete mature pept
            cds.
ACCESSION   K03194
VERSION     K03194.1  GI:182840
KEYWORDS    blood coagulation factor; factor X.
SOURCE      Human adult liver, cDNA to mRNA, clones pcHX[5,8,14].
```

```
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2)
  AUTHORS   MacGillivray,R.T.A.
  JOURNAL   Unpublished (1985)
REFERENCE   2  (bases 3 to 1443)
  AUTHORS   Fung,M.R., Hay,C.W. and MacGillivray,R.T.
  TITLE     Characterization of an almost full-length cDNA coding for human
            blood coagulation factor X
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 82 (11), 3591-3595 (1985)
  MEDLINE   85216545
   PUBMED   2582420
COMMENT     During conversion of factor X to factor X-a, a glycopeptide of 52
            amino acids (encoded by positions 513-668 in this sequence) is
            released.
            A polyadenylation signal is located at position 1424-1429. This
            sequence was kindly submitted over electronic mail by
            R.T.A.MacGillivray (23-SEP-1985).
FEATURES             Location/Qualifiers
     source          1. .1443
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /map="13q34"
     gene            1. .1443
                     /gene="F10"
     mRNA            <1. .1443
                     /gene="F10"
                     /product="X mRNA"
     CDS             1. .1433
                     /partial
                     /gene="F10"
                     /note="factor X prepeptide"
                     /codon_start=3
                     /protein_id="AAA52490.1"
                     /db_xref="GI:182841"
                     /db_xref="GDB:G00-119-890"
                     /translation="SLAGLLLLGESLFIRREQANNILARVTRANSFLEEMKKGHLERE
                     CMEETCSYEEAREVFEDSDKTNEFWNKYK GDQCETSPCQNQGKCKDGLGEYTCTCLE
                     GFEGKNCELFTRKLCSLDNGDCDQFCHEEQNSVVCSCARGYTLADNGKACIPTGPYPC
                     GKQTLERRKRSVAQATSSSGEAPDSITWKPYDAADLDPTENPFDLLDFNQTQPERGDN
                     NLTRIVGGQECKDGECPWQALLINEENEGFCGGTILSEFYILTAAHCLYQAKRFKVRV
                     GDRNTEQEEGGEAVHEVEVVIKHNRFTKETYDFDIAVLRLKTPITFRMNVAPACLPER
                     DWAESTLMTQKTGIVSGFGRTHEKGRQSTRLKMLEVPYVDRNSCKLSSSFIITQNMFC
                     AGYDTKQEDACQGDSGGPHVTRFKDTYFVTGIVSWGESCARKGKYGIYTKVTAFLKWI
                     DRSMKTRGLPKAKSHAPEVITSSPLK"
     sig_peptide     <1. .86
                     /gene="F10"
                     /note="factor X signal peptide"
     mat_peptide     87. .503
                     /gene="F10"
                     /product="factor X light chain"
     mat_peptide     513. .1430
                     /gene="F10"
                     /product="factor X heavy chain"
     variation       1117
                     /gene="F10"
                     /note="a in pcHX8; t in pcHX5"
BASE COUNT      361 a    416 c    435 g    231 t
ORIGIN      24 bp upstream of AvaI site.

  Query Match             91.0%;  Score 91;  DB 9;  Length 1443;
  Best Local Similarity   94.9%;  Pred. No. 3.9e-21;
  Matches   94;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       2 GTCTGTCACAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 61
           |||  | ||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db     703 GTCCCTGGCAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACTATTC 762

Qy      62 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 100
           |||||||||||||||||||||||||||||||||||||||
Db     763 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 801
```

```
RESULT 14
A86859
LOCUS       A86859                  1467 bp    DNA     linear   PAT 22-JAN-2000
DEFINITION  Sequence 43 from Patent WO9838318.
ACCESSION   A86859
VERSION     A86859.1  GI:6735650
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 1467)
  AUTHORS   Falkner,F. and Himmelspach,M.
  TITLE     FACTOR X DELETION MUTANTS AND ANALOGUES THEREOF
  JOURNAL   Patent: WO 9838318-A 43 03-SEP-1998;
            FALKNER FALKO GUENTER (AT); HIMMELSPACH MICHELE (AT)
FEATURES             Location/Qualifiers
     source          1. .1467
                     /organism="unidentified"
                     /db_xref="taxon:32644"
     CDS             1. .1467
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAB69367.1"
                     /db_xref="GI:6735651"
                     /translation="MGRPLHLVLLSASLAGLLLLGESLFIRREQANNILARVTRANSF
                     LEEMKKGHLERECMEETCSYEEAREVFEDSDKTNEFWNKYKDGDQCETSPCQNQGKCK
                     DGLGEYTCTCLEGFEGKNCELFTRKLCSLDNGDCDQFCHEEQNSVVCSCARGYTLADN
                     GKACIPTGPYPCGKQTLERRKRSVAQATSSSGEAPDSITWKPYDAADLDPTENPFDLL
                     DFNQTQPERGDNNLTRIVGGQECKDGECPWQALLINEENEGFCGGTILSEFYILTAAH
                     CLYQAKRFKVRVGDRNTEQEEGGEAVHEVEVVIKHNRFTKETYDFDIAVLRLKTPITF
                     RMNVAPACLPERDWAESTLMTQKTGIVSGFGRTHEKGRQSTRLKMLEVPYVDRNSCKL
                     SSSFIITQNMFCAGYDTKQEDACQGDSGGPHVTRFKDTYFVTGIVSWGESCARKGKYG
                     IYTKVTAFLKWIDRSMKTRGLPKAKSHAPEVITSSPLK"
BASE COUNT      363 a    424 c    444 g    236 t
ORIGIN

  Query Match             91.0%;  Score 91;  DB 6;  Length 1467;
  Best Local Similarity   94.9%;  Pred. No. 3.9e-21;
  Matches   94;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       2 GTCTGTCACAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 61
           |||  | ||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db     737 GTCCCTGGCAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACTATTC 796

Qy      62 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 100
           |||||||||||||||||||||||||||||||||||||||
Db     797 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 835
```

```
RESULT 15
A86886
LOCUS       A86886                  1467 bp    DNA     linear   PAT 22-JAN-2000
DEFINITION  Sequence 26 from Patent WO9838317.
ACCESSION   A86886
VERSION     A86886.1  GI:6735677
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 1467)
  AUTHORS   Himmelspach,M. and Eibl,J.
  TITLE     FACTOR X ANALOGUES WITH A MODIFIED PROTEASE CLEAVAGE SITE
  JOURNAL   Patent: WO 9838317-A 26 03-SEP-1998;
            HIMMELSPACH MICHELE (AT); EIBL JOHANN (AT)
FEATURES             Location/Qualifiers
     source          1. .1467
                     /organism="unidentified"
                     /db_xref="taxon:32644"
     CDS             1. .1467
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAB69368.1"
```

```
                     /db_xref="GI:6735678"
                     /translation="MGRPLHLVLLSASLAGLLLLGESLFIRREQANNILARVTRANSF
                     LEEMKKGHLERECMEETCSYEEAREVFEDSDKTNEFWNKYKDGDQCETSPCQNQGKCK
                     DGLGEYTCTCLEGFEGKNCELFTRKLCSLDNGDCDQFCHEEQNSVVCSCARGYTLADN
                     GKACIPTGPYPCGKQTLERRKRSVAQATSSSGEAPDSITWKPYDAADLDPTENPFDLL
                     DFNQTQPERGDNNLTRIVGGQECKDGECPWQALLINEENEGFCGGTILSEFYILTAAH
                     CLYQAKRFKVRVGDRNTEQEEGGEAVHEVEVVIKHNRFTKETYDFDIAVLRLKTPITF
                     RMNVAPACLPERDWAESTLMTQKTGIVSGFGRTHEKGRQSTRLKMLEVPYVDRNSCKL
                     SSSFIITQNMFCAGYDTKQEDACQGDSGGPHVTRFKDTYFVTGIVSWGESCARKGKYG
                     IYTKVTAFLKWIDRSMKTRGLPKAKSHAPEVITSSPLK"
BASE COUNT      363 a    424 c    444 g    236 t
ORIGIN

  Query Match           91.0%;  Score 91;  DB 6;  Length 1467;
  Best Local Similarity  94.9%;  Pred. No. 3.9e-21;
  Matches   94;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       2 GTCTGTCACAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 61
           |||  |  ||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db     737 GTCCCTGGCAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACTATTC 796

Qy      62 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 100
           ||||||||||||||||||||||||||||||||||||||||
Db     797 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 835


Search completed: January 15, 2003, 20:35:36
Job time : 1354 secs
```

ORIGINAL
COPY

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: January 15, 2003, 17:50:20 ; Search time 151 Seconds
(without alignments)
1491.390 Million cell updates/sec

Title: L00396_COPY_1_100
Perfect score: 100
Sequence: 1 CGTCTGTCACAGGCCCTGCT.....AACGGCAGCCCACTGTCTCT 100

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	100	100.0	236	21	AAC70863	Single nucleotide
	2	99.6	99.6	236	21	AAC70866	Single nucleotide
	3	92.6	92.6	725	20	AAX78022	Chimeric serine pr
	4	92.6	92.6	1126	20	AAX15427	DNA encoding coagu
	5	92.6	92.6	1126	21	AAA89786	DNA encoding coagu
	6	92.6	92.6	1126	21	AAA12970	DNA encoding Facto
	7	92.6	92.6	1126	21	AAZ56120	Vitamin-K-dependen
	8	92.6	92.6	1404	19	AAV10462	Human Factor X pro
	9	92.6	92.6	1560	22	AAF24735	Nucleotide sequenc
	10	92.6	92.6	1860	22	AAF24738	Nucleotide sequenc
	11	91	91.0	1467	19	AAV56776	Human Factor X gen
	12	91	91.0	1467	19	AAV56821	Human Factor X gen
	13	91	91.0	1467	22	AAF59409	Human factor X nuc
	14	91	91.0	1507	21	AAA54031	Human factor X cod
	15	91	91.0	1887	22	AAH57469	Human liver cell s
c	16	75.2	75.2	591	24	ABK30368	Human G-protein-co
	17	63.4	63.4	1554	15	AAQ71243	Serine protease fo
	18	36	36.0	1028	21	AAA61561	cDNA encoding mous
	19	34.4	34.4	1250	23	ABL12391	Drosophila melanog
c	20	34.4	34.4	3250	23	ABL12390	Drosophila melanog
c	21	34.4	34.4	9594	23	ABL13694	Drosophila melanog
	22	34.2	34.2	699	18	AAT79127	Human serine prote
	23	34.2	34.2	1189	24	AAI64290	Protease D-G catay
	24	34.2	34.2	1305	22	AAD02556	Human seripancrin
	25	34.2	34.2	1479	22	AAD02557	Human seripancrin
	26	34.2	34.2	1795	22	AAS26880	Human cDNA encodin
c	27	34.2	34.2	1854	22	AAH99574	Human protein enco
	28	34.2	34.2	2038	20	AAX87154	Human protease HUP
	29	34.2	34.2	2063	21	AAA37099	Human PRO1570 (UNQ
	30	34.2	34.2	2063	22	AAS46089	Human DNA encoding
	31	34.2	34.2	2063	22	AAF92113	Human PRO1570 cDNA
	32	34.2	34.2	2063	22	AAF54396	DNA encoding prote
	33	34.2	34.2	2063	24	ABK11090	cDNA encoding tumo
	34	34.2	34.2	2079	22	AAD13169	Human transmembran
	35	34.2	34.2	2079	24	AAI72975	CJA8 cDNA. Homo s
	36	34.2	34.2	2081	24	AAI72976	CJA8 preferred cDN
	37	34.2	34.2	2121	24	AAI64284	Human serine prote
	38	33.8	33.8	804	23	ABL28901	Drosophila melanog
	39	33.8	33.8	2864	23	ABL28900	Drosophila melanog
	40	32.8	32.8	180	21	AAZ89010	Cavea sp. Factor I
	41	32.6	32.6	2137	22	AAD13114	Human membrane-typ
	42	32	32.0	817	23	ABL04003	Drosophila melanog
	43	32	32.0	846	23	ABL17189	Drosophila melanog
	44	32	32.0	864	19	AAV04616	Flea serine protea
	45	32	32.0	864	22	AAC90884	Flea serine protea

ALIGNMENTS

RESULT 1
AAC70863
ID AAC70863 standard; DNA; 236 BP.
XX
AC AAC70863;
XX
DT 09-FEB-2001 (first entry)
XX
DE Single nucleotide polymorphism containing sequence #231.
XX
KW Single nucleotide polymorphism; SNP; human; genetic disease;
KW disease susceptibility; cardiovascular system; endocrine system;
KW neurological system; forensic testing; paternity testing; ds.
XX
OS Homo sapiens.
XX
PN WO200058519-A2.
XX
PD 05-OCT-2000.
XX
PF 30-MAR-2000; 2000WO-US08440.
XX
PR 31-MAR-1999; 99US-0127248.
XX
PA (WHED) WHITEHEAD INST BIOMEDICAL RES.
PA (AFFY-) AFFYMETRIX INC.
XX
PI Altshuler D, Cargill M, Daley GQ, Ireland JS, Lander ES;
PI Lipshutz RJ, Patil N, Sklar P;
XX
DR WPI; 2000-611722/58.
XX

```
PT   Nucleic acid selected from one of 106 genes comprising single
PT   nucleotide polymorphisms, allele-specific oligonucleotides to the genes
PT   are useful for phenotypic correlations, forensics, paternity testing,
PT   medicine and genetic analysis -
XX
PS   Claim 1; Fig 5; 214pp; English.
XX
CC   The present invention is concerned with a number of human single
CC   nucleotide polymorphisms (SNPs) which the inventors identified in human
CC   genes. These SNPs can be used in disease diagnosis and prediction of an
CC   individual's susceptibility to disease, in forensic and paternity testing
CC   and in genetic mapping. In particular, the SNPs of the invention can be
CC   used to diagnose susceptibility to diseases of the cardiovascular,
CC   endocrine and neurological systems, such as coronary artery disease,
CC   schizophrenia, cancer, autoimmune diseases, Alzheimer's and Parkinson's
CC   diseases.
CC   Note: The degenerate codon within the sequence represents the position
CC   of an SNP, for example the letter S represents a polymorphism where the
CC   nucleotide may be C or G.
XX
SQ   Sequence 236 BP; 54 A; 66 C; 67 G; 47 T; 2 other;

  Query Match             100.0%;  Score 100;  DB 21;  Length 236;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-25;
  Matches  100;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CGTCTGTCACAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      54 CGTCTGTCACAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATT 113

Qy      61 CTGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 100
           |||||||||||||||||||||||||||||||||||||||
Db     114 CTGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 153


RESULT 2
AAC70866
ID   AAC70866 standard; DNA; 236 BP.
XX
AC   AAC70866;
XX
DT   09-FEB-2001  (first entry)
XX
DE   Single nucleotide polymorphism containing sequence #232.
XX
KW   Single nucleotide polymorphism; SNP; human; genetic disease;
KW   disease susceptibility; cardiovascular system; endocrine system;
KW   neurological system; forensic testing; paternity testing; ds.
XX
OS   Homo sapiens.
XX
PN   WO200058519-A2.
XX
PD   05-OCT-2000.
XX
PF   30-MAR-2000; 2000WO-US08440.
XX
PR   31-MAR-1999;   99US-0127248.
XX
PA   (WHED ) WHITEHEAD INST BIOMEDICAL RES.
PA   (AFFY-) AFFYMETRIX INC.
XX
PI   Altshuler D,  Cargill M,  Daley GQ,  Ireland JS,  Lander ES;
PI   Lipshutz RJ,  Patil N,  Sklar P;
XX
DR   WPI; 2000-611722/58.
XX
PT   Nucleic acid selected from one of 106 genes comprising single
PT   nucleotide polymorphisms, allele-specific oligonucleotides to the genes
PT   are useful for phenotypic correlations, forensics, paternity testing,
PT   medicine and genetic analysis -
XX
PS   Claim 1; Fig 5; 214pp; English.
XX
CC   The present invention is concerned with a number of human single
CC   nucleotide polymorphisms (SNPs) which the inventors identified in human
CC   genes. These SNPs can be used in disease diagnosis and prediction of an
CC   individual's susceptibility to disease, in forensic and paternity testing
CC   and in genetic mapping. In particular, the SNPs of the invention can be
CC   used to diagnose susceptibility to diseases of the cardiovascular,
CC   endocrine and neurological systems, such as coronary artery disease,
CC   schizophrenia, cancer, autoimmune diseases, Alzheimer's and Parkinson's
CC   diseases.
CC   Note: The degenerate codon within the sequence represents the position
CC   of an SNP, for example the letter S represents a polymorphism where the
CC   nucleotide may be C or G.
XX
SQ   Sequence 236 BP; 54 A; 66 C; 67 G; 47 T; 2 other;

  Query Match             99.6%;  Score 99.6;  DB 21;  Length 236;
  Best Local Similarity   99.0%;  Pred. No. 1.6e-25;
  Matches   99;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CGTCTGTCACAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||
Db      54 CGTCTGTCACAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACYATT 113

Qy      61 CTGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 100
           |||||||||||||||||||||||||||||||||||||||
Db     114 CTGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 153


RESULT 3
AAX78022
ID   AAX78022 standard; DNA; 725 BP.
XX
AC   AAX78022;
XX
DT   19-AUG-1999  (first entry)
XX
DE   Chimeric serine protease FXT DNA.
XX
KW   Serine protease; chimeric; antithrombotic; modulator; drug design;
KW   3-D crystal structure; crystallization; haematopoietic cascade;
KW   FXT; ss.
XX
OS   Synthetic.
XX
PN   EP927764-A2.
XX
PD   07-JUL-1999.
XX
PF   27-NOV-1998;   98EP-0122481.
XX
PR   03-DEC-1997;   97EP-0121232.
XX
PA   (HOFF ) ROCHE DIAGNOSTICS GMBH.
XX
PI   Bode W,  Engh R,  Hopfner K,  Huber R,  Kopetzki E;
XX
DR   WPI; 1999-359878/31.
DR   P-PSDB; AAY08894.
XX
PT   Chimeric serine protease comprising Factor X and Trypsin catalytic
PT   domains, useful for identifying thrombolytic agents
XX
PS   Example 2; Fig 2; 23pp; German.
XX
CC   This invention describes a novel chimeric serine protease compound which
CC   has antithrombotic activity and comprises two domains with a beta-sheet
CC   structure. The first domain corresponds to the first domain of a first
CC   serine protease and the second domain corresponds to the second domain
CC   of a second serine protease. The products of the invention can be used
CC   for identifying antithrombotic agents by determining whether an agent
CC   modulates the activity of the serine protease. They can also be used to
CC   identify agents through rational drug design using information based on
```

```
CC   its 3-D crystal structure. The chimeric serine protease, is very good
CC   for crystallization and for determining structural data and is also
CC   useful for the identification of specific antithrombotic agents that
CC   unlike prior art agents are extremely specific for only one factor in a
CC   haematopoietic cascade.
XX
SQ   Sequence 725 BP; 172 A; 198 C; 216 G; 139 T; 0 other;

  Query Match             92.6%;  Score 92.6;  DB 20;  Length 725;
  Best Local Similarity   96.0%;  Pred. No. 6.6e-23;
  Matches   95;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 GTCTGTCACAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 61
           |||  |  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      65 GTCCCTGGCAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 124

Qy      62 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 100
           ||||||||||||||||||||||||||||||||||||||||
Db     125 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 163


RESULT 4
AAX15427
ID   AAX15427 standard; DNA; 1126 BP.
XX
AC   AAX15427;
XX
DT   05-MAY-1999  (first entry)
XX
DE   DNA encoding coagulation factor X/Xa.
XX
KW   Truncated tissue factor; tissue factor binding ligand; coagulation;
KW   disease-associated vasculature; tumour; benign prostatic hyperplasia;
KW   diabetic-retinopathy; vascular restenosis; arteriovenous malformation;
KW   AVM; meningioma; hemangioma; neovascular glaucoma; psoriasis; synovitis;
KW   dermatitis; endometriosis; angiofibroma; rheumatoid arthritis;
KW   atheroscelotic plaque; corneal graft neovascularisation;
KW   haemophilic joint; hypertrophic scar; Osler-Weber syndrome;
KW   pyogenic granuloma retrolental fibroplasia; scleroderma; trachoma;
KW   vascular adhesion; coagulation factor; factor X/Xa; ss.
XX
OS   Homo sapiens.
XX
PN   US5877289-A.
XX
PD   02-MAR-1999.
XX
PF   07-JUN-1995;   95US-0479733.
XX
PR   07-JUN-1995;   95US-0479733.
PR   05-MAR-1992;   92US-0846349.
PR   02-MAR-1994;   94US-0205330.
PR   11-JUL-1994;   94US-0273567.
XX
PA   (SCRI ) SCRIPPS RES INST.
PA   (TEXA ) UNIV TEXAS SYSTEM.
XX
PI   Edgington TS,  Thorpe PE;
XX
DR   WPI; 1999-189722/16.
XX
PT   Tissue factor binding ligands  - comprising first binding region
PT   which binds to vasculature, particularly of tumours, and tissue
PT   factor construct
XX
PS   Example 9; Columns 129-132; 83pp; English.
XX
CC   The present sequence encodes a coagulation factor. The specification
CC   describes tissue factor binding ligands which comprise a binding region
CC   which binds to vasculature, particularly of tumours, and a tissue factor
CC   construct. The binding ligands can be used for stimulating coagulation in
CC   disease-associated vasculature, particularly for the treatment of
CC   tumours. The products can also be used for treating e.g. benign prostatic
CC   hyperplasia, diabetic-retinopathy, vascular restenosis, arteriovenous
CC   malformations (AVM), meningioma, hemangioma, neovascular glaucoma,
CC   psoriasis, synovitis, dermatitis, endometriosis, angiofibroma, rheumatoid
CC   arthritis, atheroscelotic plaques, corneal graft neovascularisation,
CC   haemophilic joints, hypertrophic scars, Osler-Weber syndrome, pyogenic
CC   granuloma retrolental fibroplasia, scleroderma, trachoma, or vascular
CC   adhesions. The products can also be used in binding assays.
XX
SQ   Sequence 1126 BP; 269 A; 341 C; 342 G; 174 T; 0 other;

  Query Match             92.6%;  Score 92.6;  DB 20;  Length 1126;
  Best Local Similarity   96.0%;  Pred. No. 7.5e-23;
  Matches   95;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 GTCTGTCACAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 61
           |||  |  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     395 GTCCCTGGCAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 454

Qy      62 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 100
           ||||||||||||||||||||||||||||||||||||||||
Db     455 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 493


RESULT 5
AAA89786
ID   AAA89786 standard; DNA; 1126 BP.
XX
AC   AAA89786;
XX
DT   14-DEC-2000  (first entry)
XX
DE   DNA encoding coagulation factor X/Xa.
XX
KW   Tissue factor protein; truncated tissue factor; tTF; cytostatic;
KW   coagulant; diabetic retinopathy; arteriovenous malformation;
KW   meningioma; hemangioma; neovascular glaucoma; psoriasis; synovitis;
KW   endometriosis; hemophylic joint; hypertrophic scar; vascular adhesion;
KW   tumour; cancer; ligand; human; factor X; ds.
XX
OS   Homo sapiens.
XX
PN   US6093399-A.
XX
PD   25-JUL-2000.
XX
PF   07-JUN-1995;   95US-0482369.
XX
PR   05-MAR-1992;   92US-0846349.
PR   02-MAR-1994;   94US-0205330.
PR   11-JUL-1994;   94US-0273567.
XX
PA   (SCRI ) SCRIPPS RES INST.
PA   (TEXA ) UNIV TEXAS SYSTEM.
XX
PI   Edgington TS,  Thorpe PE;
XX
DR   WPI; 2000-531471/48.
XX
PT   New immunological and growth factor-based bispecific binding ligands,
PT   useful for stimulating coagulation in vasculature-associated diseases,
PT   e.g. for treating both benign and malignant diseases (e.g. meningioma
PT   or hemangioma)
XX
PS   Example 9; Column 129-130; 83pp; English.
XX
CC   The present invention relates to a binding ligand with a first binding
CC   region that is operatively linked to either a coagulation factor or a
CC   second binding region that binds to a coagulation factor. The first
CC   binding region binds to a component on the surface of a tumour. The
CC   second binding region is all or part of an antibody. An example of a
CC   coagulation factor for use in the invention is human truncated tissue
CC   factor. Truncated tissue factor (tTF) is the extracellular domain of the
CC   mature tissue factor protein (see AAB15019). The binding ligand of the
```

```
CC   invention is useful for stimulating coagulation in vasculature
CC   associated diseases. Particularly, the binding ligand is useful for
CC   treating both benign and malignant diseases that have a vascular
CC   component. These diseases include benign growths (e.g. BPH), diabetic
CC   retinopathy, arteriovenous malformations, meningioma, hemangioma,
CC   neovascular glaucoma, psoriasis, synovitis, endometriosis, hemophylic
CC   joints, hypertrophic scars or vascular adhesions. The present binding
CC   ligands offer the advantage that even limited damage to the tumour
CC   vasculature could produce an avalanche of tumour cell death because
CC   each capillary provides oxygen and nutrients for thousands of tumour
CC   cells. The present sequence is DNA encoding coagulation factor
CC   X/Xa. This factor was used in the invention.
XX
SQ   Sequence 1126 BP; 269 A; 341 C; 342 G; 174 T; 0 other;

  Query Match             92.6%;  Score 92.6;  DB 21;  Length 1126;
  Best Local Similarity   96.0%;  Pred. No. 7.5e-23;
  Matches   95;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 GTCTGTCACAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 61
           |||  |  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     395 GTCCCTGGCAGGCCCTGCTCATCAATGAGGAAAACGAC GCTTTCTGTGGTGGAACCAT C 454

Qy      62 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 100
           |||||||||||||||||||||||||||||||||||||||
Db     455 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 493


RESULT 6
AAA12970
ID   AAA12970 standard; DNA; 1126 BP.
XX
AC   AAA12970;
XX
DT   18-JUL-2000  (first entry)
XX
DE   DNA encoding Factor X/Xa, SEQ ID NO:27.
XX
KW   Truncated tissue factor; tTF; human; blood coagulation;
KW   tumour vasculature; bispecific antibody; targetting; cancer;
KW   vascularised tumour; PCR primer; ss.
XX
OS   Homo sapiens.
XX
PN   US6036955-A.
XX
PD   14-MAR-2000.
XX
PF   07-JUN-1995;   95US-0479727.
XX
PR   05-MAR-1992;   92US-0846349.
PR   02-MAR-1994;   94US-0205330.
PR   11-JUL-1994;   94US-0273567.
XX
PA   (TEXA ) UNIV TEXAS SYSTEM.
PA   (SCRI ) SCRIPPS RES INST.
XX
PI   Edgington TS,  Thorpe PE;
XX
DR   WPI; 2000-269871/23.
XX
PT   Kit for inducing coagulation in tumor vasculature, useful for treating
PT   malignant or benign growths, contains ligand, linked to coagulation
PT   agent, that targets tumor marker -
XX
PS   Example 9; Columns 131-132; 86pp; English.
XX
CC   The invention relates to the induction of blood coagulation specifically
CC   within tumour vasculature. This is achieved by the use of a bispecific
CC   molecule, which comprises a region capable of binding to intratumoral
CC   vascular or stromal cells linked to a coagulation factor or to a region
CC   capable of binding to a coagulation factor. An example of such a
CC   bispecific molecule is a bispecific antibody, where one arm binds a
CC   tumour antigen, and the other arm binds a coagulation factor. The
CC   expression of certain proteins (tumour antigens) is upregulated in
CC   tumour vasculature; such proteins include vascular endothelial growth
CC   factor (VEGF) and members of the fibroblast growth factor (FGF) family.
CC   An antibody or antibody fragment against VEGF or basic FGF (bFGF) may be
CC   incorporated into the bispecific molecule in order to target coagulation
CC   to tumour vasculature. The coagulation factor-binding portion of the
CC   bispecific molecule may be, for example, directed to tissue factor (TF).
CC   A preferred form of TF used in the invention is a truncated form (tTF,
CC   AAY81488) which lacks the cytoplasmic and transmembrane domains.
CC   Although tTF can associate with Factor VIIa, the tTF/Factor VIIa complex
CC   cannot alone initiate the coagulation cascade as the complex has to be
CC   associated with a phospholipid surface for coagulation to occur.
CC   However, binding of tTF to tumour vasculature via a tumour antigen/tTF
CC   bispecific antibody brings tTF into close enough proximity with the cell
CC   membrane to enable the initiation of coagulation. Kits for the induction
CC   of tumour vasculature-specific coagulation may be used to treat malignant
CC   or benign diseases associated with a vascular component, particularly
CC   cancers, but also benign growths, prostatic hypertrophy, restenosis,
CC   psoriasis, glaucoma, rheumatoid arthritis. Coagulation is induced
CC   selectively in the tumour vasculature, minimising side effects. Such kits
CC   are likely to be effective against many different types of cancer.
CC   Sequences AAA12945- AAA12952, AAA12954-A12963 and AAA12971-A12972
CC   represent PCR primers used in exemplifications of the present invention
CC   to generate constructs encoding tTF, tTF variants or tTF dimers.
XX
SQ   Sequence 1126 BP; 269 A; 341 C; 342 G; 174 T; 0 other;

  Query Match             92.6%;  Score 92.6;  DB 21;  Length 1126;
  Best Local Similarity   96.0%;  Pred. No. 7.5e-23;
  Matches   95;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 GTCTGTCACAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 61
           |||  |  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     395 GTCCCTGGCAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 454

Qy      62 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 100
           |||||||||||||||||||||||||||||||||||||||
Db     455 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 493


RESULT 7
AAZ56120
ID   AAZ56120 standard; DNA; 1126 BP.
XX
AC   AAZ56120;
XX
DT   27-MAR-2000  (first entry)
XX
DE   Vitamin-K-dependent coagulation factor X/Xa coding sequence.
XX
KW   Vitamin-K dependent coagulation factor; tumour associated vasculature;
KW   carcinoma; benign prostatic hyperplasia; diabetic retinopathy;
KW   vascular restenosis; arteriovenous malformation; meningoma; haemangioma;
KW   neovascular glaucoma; psoriasis; cytostatic; antidiabetic; vasotropic;
KW   ophthalmological; antipsoriatic; Factor X/Xa; ss.
XX
OS   Unspecified.
XX
PN   US6004555-A.
XX
PD   21-DEC-1999.
XX
PF   07-JUN-1995;   95US-0487427.
XX
PR   05-MAR-1992;   92US-0846349.
PR   02-MAR-1994;   94US-0205330.
PR   11-JUL-1994;   94US-0273567.
XX
PA   (SCRI ) SCRIPPS RES INST.
PA   (TEXA ) UNIV TEXAS SYSTEM.
XX
PI   Edgington TS,  Thorpe PE;
```

```
XX
DR   WPI; 2000-072047/06.
XX
PT   Bispecific binding ligands for promoting blood coagulation in a tumour
PT   associated vasculature are useful for treating cancer -
XX
PS   Example 9; Column 131-132; 83pp; English.
XX
CC   This is the coding sequence for Factor X/Xa, a vitamin-K-dependent
CC   coagulation factor. This coagulation factor can be used in the formation
CC   of coaguligands. Mutated versions of this sequence can be used in the
CC   method for delivering a coagulant to a tumour-associated vasculature
CC   using bispecific binding ligands which promote blood coagulation. The
CC   binding ligand consists of a binding region that binds to a
CC   surface-expressed, surface accessible or surface-localised component of a
CC   tumour cell, intratumoural vasculature or tumour associated stroma. The
CC   binding region is linked to a coagulating agent which is a coagulation
CC   factor (e.g. tissue factor). The second binding region comprises an
CC   antibody or an antigen binding region of an antibody. The method is used
CC   for delivering an exogenous or an endogenous coagulation factor to
CC   tumour-associated vasculature which is benign or malignant. The method
CC   can be used to treat cancer by promoting specific blood coagulation in
CC   the vasculature of the tumour relative to the vasculature in nontumour
CC   sites. Vascularised tumours are usually solid tumours, particularly
CC   carcinomas which require a vascular component to provide oxygen and
CC   nutrients. The ligands are suitable to treat benign and malignant
CC   diseases with a vascular component, including benign prostatic
CC   hyperplasia, diabetic retinopathy, vascular restenosis, arteriovenous
CC   malformations, meningioma, haemangioma, neovascular glaucoma and
CC   psoriasis. The ligands can also be used in standard binding assays in
CC   vitro. Bispecific ligands can be designed which are capable of binding to
CC   vascular endothelial cells and disease-associated agents such as
CC   activated platelets. Certain disease-associated agents are similar in
CC   different diseases and in different tumours, making it possible to treat
CC   numerous diseases and different types of cancer with one pharmaceutical,
CC   therefore an agent does not need to be tailored to each individual
CC   disease or specific tumour type.
XX
SQ   Sequence 1126 BP; 269 A; 341 C; 342 G; 174 T; 0 other;

  Query Match          92.6%;  Score 92.6;  DB 21;  Length 1126;
  Best Local Similarity   96.0%;  Pred. No. 7.5e-23;
  Matches   95;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 GTCTGTCACAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 61
           |||  |  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     395 GTCCCTGGCAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 454

Qy      62 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 100
           ||||||||||||||||||||||||||||||||||||||||
Db     455 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 493


RESULT 8
AAV10462
ID   AAV10462 standard; DNA; 1404 BP.
XX
AC   AAV10462;
XX
DT   16-JUN-1998  (first entry)
XX
DE   Human Factor X protease cDNA.
XX
KW   Factor X; factor IX; serine protease activity; catalytic domain; ZAD;
KW   zymogen-activating domain; epidermal growth factor-like domain; EGF1;
KW   EGF2; regulator; coagulation; fibrinolysis; homeostasis; X-ray structure;
KW   detection; drug modelling; restriction protease; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1404
FT                   /*tag=  a
FT                   /product= Factor X
FT                   /note= "partial coding sequence"
XX
PN   WO9747737-A1.
XX
PD   18-DEC-1997.
XX
PF   11-JUN-1997;   97WO-EP03027.
XX
PR   06-JUL-1996;   96EP-0110959.
PR   11-JUN-1996;   96EP-0109288.
PR   22-JUN-1996;   96EP-0110109.
XX
PA   (BOEF ) BOEHRINGER MANNHEIM GMBH.
XX
PI   Hopfner K,  Kopetzki E;
XX
DR   WPI; 1998-052304/05.
DR   P-PSDB; AAW40283.
XX
PT   Non-glycosylated, truncated forms of factor IX family protein with
PT   serine protease activity - used to screen for specific modulators
PT   and to assay factor IXa
XX
PS   Disclosure; Fig 3; 49pp; German.
XX
CC   This sequence encodes a human factor X protease. This protein is used
CC   in the construction of a novel non-glycosylated protein and truncated
CC   and zymogen forms of this protein, which have serine protease activity.
CC   The protein is composed of various domains from a factor IX family
CC   protein, namely a catalytic domain (CD) N-terminally bound to a
CC   zymogen-activating domain (ZAD), N-terminally bound to an EGF1 and/or
CC   EGF2 domain (EGF = epidermal growth factor-like domain). Such proteins
CC   are used to identify activators/inhibitors of factor IX family proteins
CC   (potentially useful as regulators of coagulation, fibrinolysis and
CC   homeostasis). The protein in zymogen form is also useful in assays for
CC   detecting factor IXa activity in aqueous solution (specifically in body
CC   fluids). The protein can be used to produce co-crystals with protease
CC   variants or inhibitors for X-ray structural analysis and drug modelling
CC   and as restriction proteases in biotechnology. These truncated proteins
CC   have the same specificity as factor IX family proteases and can be
CC   produced in prokaryotes in a form that allows production of active enzyme
CC   by conversion to native form and enzymatic cleavage.
XX
SQ   Sequence 1404 BP; 356 A; 404 C; 423 G; 221 T; 0 other;

  Query Match          92.6%;  Score 92.6;  DB 19;  Length 1404;
  Best Local Similarity   96.0%;  Pred. No. 8.1e-23;
  Matches   95;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 GTCTGTCACAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 61
           |||  |  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     683 GTCCCTGGCAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 742

Qy      62 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 100
           ||||||||||||||||||||||||||||||||||||||||
Db     743 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 781


RESULT 9
AAF24735
ID   AAF24735 standard; DNA; 1560 BP.
XX
AC   AAF24735;
XX
DT   20-APR-2001  (first entry)
XX
DE   Nucleotide sequence of Sig/CBD cex/ER retaining peptide fusion.
XX
KW   Protein production; food processing; protein antibiotic; feed enzyme;
KW   protein L; CBD cex protein; cell signal peptide; ss.
XX
OS   Synthetic.
```

```
XX
FH   Key             Location/Qualifiers
FT   CDS             20..1549
FT                   /*tag=  a
XX
PN   WO200077174-A1.
XX
PD   21-DEC-2000.
XX
PF   07-JUN-2000; 2000WO-IL00330.
XX
PR   10-JUN-1999;   99US-0329234.
XX
PA   (CBDT-) CBD TECHNOLOGIES LTD.
PA   (YISS ) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
XX
PI   Shani Z,  Shoseyov O;
XX
DR   WPI; 2001-112219/12.
DR   P-PSDB; AAB31373.
XX
PT   Expressing and isolating recombinant protein in a plant, useful for
PT   producing large quantities of recombinant proteins, by expressing a
PT   fusion protein including a cellulose binding peptide fused to a
PT   recombinant protein  -
XX
PS   Example; Fig 3a; 87pp; English.
XX
CC   The specification describes a method for expressing and isolating
CC   a recombinant protein in a plant. The method comprising expressing a
CC   fusion protein including the recombinant protein and a cellulose
CC   binding peptide fused to it, where the fusion protein is
CC   compartmentalised and sequestered within plant cells, plant derived
CC   tissue or cultured plant cells. The method is useful for obtaining large
CC   quantities of the recombinant proteins and protein products in a simple
CC   and cost-effective manner. Recombinant proteins may be used commercially,
CC   such as in the food processing industry, e.g. glucoamylases and glucose
CC   isomerases are used for converting starch to high fructose corn syrup,
CC   proteinases for the hydrolysis of high molecular weight proteins and in
CC   manufacturing leather or alcoholic beverages, pectinesterases for
CC   pectin hydrolysis in food industry, lipases for cleaving ester linkage
CC   in triglycerides, and for effluent treatment. The recombinant proteins
CC   may further be used to produce protein antibiotics, which can be used
CC   in healing processes, and to produce animal feed enzymes. The present
CC   sequence encodes a fusion protein of the invention, and comprises a
CC   fusion of a cell signal peptide, Sig (not specified), CBD cex, and an
CC   endoplasmic reticulum retaining peptide.
XX
SQ   Sequence 1560 BP; 336 A; 479 C; 454 G; 291 T; 0 other;

  Query Match             92.6%;  Score 92.6;  DB 22;  Length 1560;
  Best Local Similarity   96.0%;  Pred. No. 8.3e-23;
  Matches   95;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 GTCTGTCACAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 61
           |||  |  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     837 GTCCCTGGCAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 896

Qy      62 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 100
           |||||||||||||||||||||||||||||||||||||||
Db     897 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 935


RESULT 10
AAF24738
ID   AAF24738 standard; DNA; 1860 BP.
XX
AC   AAF24738;
XX
DT   20-APR-2001  (first entry)
XX
DE   Nucleotide sequence of Sig/CBD-Tma/ER retaining peptide fusion.
XX
KW   Protein production; food processing; protein antibiotic; feed enzyme;
KW   protein L; CBD cex protein; cell signal peptide; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   CDS             20..1873
FT                   /*tag=  a
XX
PN   WO200077174-A1.
XX
PD   21-DEC-2000.
XX
PF   07-JUN-2000; 2000WO-IL00330.
XX
PR   10-JUN-1999;   99US-0329234.
XX
PA   (CBDT-) CBD TECHNOLOGIES LTD.
PA   (YISS ) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
XX
PI   Shani Z,  Shoseyov O;
XX
DR   WPI; 2001-112219/12.
DR   P-PSDB; AAB31374.
XX
PT   Expressing and isolating recombinant protein in a plant, useful for
PT   producing large quantities of recombinant proteins, by expressing a
PT   fusion protein including a cellulose binding peptide fused to a
PT   recombinant protein  -
XX
PS   Example; Fig 4a; 87pp; English.
XX
CC   The specification describes a method for expressing and isolating
CC   a recombinant protein in a plant. The method comprising expressing a
CC   fusion protein including the recombinant protein and a cellulose
CC   binding peptide fused to it, where the fusion protein is
CC   compartmentalised and sequestered within plant cells, plant derived
CC   tissue or cultured plant cells. The method is useful for obtaining large
CC   quantities of the recombinant proteins and protein products in a simple
CC   and cost-effective manner. Recombinant proteins may be used commercially,
CC   such as in the food processing industry, e.g. glucoamylases and glucose
CC   isomerases are used for converting starch to high fructose corn syrup,
CC   proteinases for the hydrolysis of high molecular weight proteins and in
CC   manufacturing leather or alcoholic beverages, pectinesterases for
CC   pectin hydrolysis in food industry, lipases for cleaving ester linkage
CC   in triglycerides, and for effluent treatment. The recombinant proteins
CC   may further be used to produce protein antibiotics, which can be used
CC   in healing processes, and to produce animal feed enzymes. The present
CC   sequence encodes a fusion protein of the invention, and comprises a
CC   fusion of a cell signal peptide, Sig (not specified), CBD-Tma, and an
CC   endoplasmic reticulum retaining peptide.
XX
SQ   Sequence 1860 BP; 476 A; 516 C; 555 G; 313 T; 0 other;

  Query Match             92.6%;  Score 92.6;  DB 22;  Length 1860;
  Best Local Similarity   96.0%;  Pred. No. 8.8e-23;
  Matches   95;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 GTCTGTCACAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 61
           |||  |  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1101 GTCCCTGGCAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 1160

Qy      62 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 100
           |||||||||||||||||||||||||||||||||||||||
Db    1161 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 1199


RESULT 11
AAV56776
ID   AAV56776 standard; DNA; 1467 BP.
XX
AC   AAV56776;
XX
```

```
DT   27-NOV-1998  (first entry)
XX
DE   Human Factor X genomic DNA.
XX
KW   Factor X; analogue; activation cleavage site; protease; bleeding; human;
KW   Factor IX; Factor VII; Factor VIII; haemophilia; gene therapy; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1467
FT                   /*tag= a
FT                   /product= "Factor X"
FT   sig_peptide     1..120
FT                   /*tag= b
FT   mat_peptide     121..1464
FT                   /*tag= c
XX
PN   WO9838317-A1.
 X
PD   03-SEP-1998.
XX
PF   27-FEB-1998;   98WO-AT00045.
XX
PR   27-FEB-1997;   97AT-0000335.
XX
PA   (IMMO ) IMMUNO AG.
XX
PI   Dorner F,  Eibl J,  Fisch A,  Himmelspach M,  Schlokat U;
XX
DR   WPI; 1998-481211/41.
DR   P-PSDB; AAW76216.
XX
PT   New factor X analogues with processing site for protease not active
PT   on natural protein - and related DNA, is very stable and can be
PT   activated in vitro or in vivo without using animal protease(s),
PT   particularly for treating disorders of blood coagulation
XX
PS   Claim 3; Fig 1; 86pp; German.
XX
CC   This sequence encodes the human Factor X protein which is used in a
CC   method resulting in the production of novel human Factor X (F10)
CC   analogues. Such analogues have in the region of the natural F10a
CC   activation cleavage site, a modification that creates a processing site
CC   for a protease that does not naturally cleave F10 in this region. The
CC   proteins are used to generate, in vivo or in vitro, F10a analogues that
CC   can be used to control bleeding and for treating defects of factors IX,
CC   VII or VIII, e.g.in haemophiliacs who have developed antibodies to
CC   factors VIII and/or IX. The encoding nucleic acid can be used in gene
CC   therapy of the same conditions. The analogues have high stability and can
CC   be activated without use of animal enzymes such as trypsin. Only
CC   activation is affected, their activity is the same as the natural factor.
CC   The analogues can be isolated as a pure single-chain pro-protein (not
CC   usually possible because of rapid processing of the native precursor) and
CC   this converted to two-chain form by subsequent activation. Activated
CC   analogues have good stability and structural integrity and are
CC   practically free of inactive intermediates and autoproteolytic
CC   decomposition products.
XX
SQ   Sequence 1467 BP; 363 A; 424 C; 444 G; 236 T; 0 other;

  Query Match            91.0%;  Score 91;  DB 19;  Length 1467;
  Best Local Similarity  94.9%;  Pred. No. 3e-22;
  Matches   94;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       2 GTCTGTCACAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 61
           |||  |  |||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db     737 GTCCCTGGCAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACTATTC 796

Qy      62 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 100
           |||||||||||||||||||||||||||||||||||||||
Db     797 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 835
```

```
RESULT 12
AAV56821
ID   AAV56821 standard; DNA; 1467 BP.
XX
AC   AAV56821;
XX
DT   27-NOV-1998  (first entry)
XX
DE   Human Factor X genomic DNA.
XX
KW   Factor X; analogue; activation cleavage site; protease; bleeding; human;
KW   Factor IX; Factor VII; Factor VIII; haemophilia; gene therapy; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1467
FT                   /*tag= a
FT   sig_peptide     1..120
FT                   /*tag= b
FT   mat_peptide     121..1464
FT                   /*tag= c
FT                   /product= "Factor X"
XX
PN   WO9838318-A1.
XX
PD   03-SEP-1998.
XX
PF   27-FEB-1998;   98WO-AT00046.
XX
PR   27-FEB-1997;   97AT-0000336.
XX
PA   (IMMO ) IMMUNO AG.
XX
PI   Dorner F,  Eibl J,  Falkner F,  Himmelspach M,  Pfleiderer M;
PI   Schlokat U;
XX
DR   WPI; 1998-481212/41.
DR   P-PSDB; AAW76218.
XX
PT   New factor 10 deletion mutants lacking the natural protease
PT   processing site - but having a non-natural site inserted, and
PT   related DNA, particularly for in vitro activation to products used
PT   to treat blood coagulation disorders
XX
PS   Claim 3; Fig 1; 82pp; German.
XX
CC   This sequence encodes the human Factor X protein which is used in a
CC   method resulting in the production of novel human Factor X (F10)
CC   analogues. Such analogues have in the region of the natural F10a
CC   activation cleavage site, a modification that creates a processing site
CC   for a protease that does not naturally cleave F10 in this region. The
CC   proteins are used to generate, in vivo or in vitro, F10a analogues that
CC   can be used to control bleeding and for treating defects of factors IX,
CC   VII or VIII, e.g.in haemophiliacs who have developed antibodies to
CC   factors VIII and/or IX. The encoding nucleic acid can be used in gene
CC   therapy of the same conditions. The analogues have high stability and can
CC   be activated without use of animal enzymes such as trypsin. Only
CC   activation is affected, their activity is the same as the natural factor.
CC   The analogues can be isolated as a pure single-chain pro-protein (not
CC   usually possible because of rapid processing of the native precursor) and
CC   this converted to two-chain form by subsequent activation. Activated
CC   analogues have good stability and structural integrity and are
CC   practically free of inactive intermediates and autoproteolytic
CC   decomposition products.
XX
SQ   Sequence 1467 BP; 363 A; 424 C; 444 G; 236 T; 0 other;

  Query Match            91.0%;  Score 91;  DB 19;  Length 1467;
  Best Local Similarity  94.9%;  Pred. No. 3e-22;
  Matches   94;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;
```

```
Qy       2 GTCTGTCACAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 61
           |||  |  ||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db     737 GTCCCTGGCAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACTATTC 796

Qy      62 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 100
           |||||||||||||||||||||||||||||||||||||||
Db     797 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 835


RESULT 13
AAF59409
ID   AAF59409 standard; cDNA; 1467 BP.
XX
AC   AAF59409;
XX
DT   02-MAY-2001  (first entry)
XX
DE   Human factor X nucleotide sequence SEQ ID NO:1.
XX
KW   Human; factor X; mutant; haemostatic; gene therapy; haemophilia;
KW   blood coagulation disorder; haemophiliac; ss.
XX
OS   Homo sapiens.
XX
PN   WO200110896-A2.
XX
PD   15-FEB-2001.
XX
PF   07-AUG-2000; 2000WO-EP07631.
XX
PR   10-AUG-1999;   99AT-0001377.
XX
PA   (BAXT ) BAXTER AG.
XX
PI   Himmelspach M, Schlokat U;
XX
DR   WPI; 2001-191516/19.
DR   P-PSDB; AAB70411.
XX
PT   Novel factor X analog useful for producing drug which is useful for
PT   treatment of blood coagulation disorders, such as hemophilia, contains
PT   modification between amino acids Glu226 and Ile235
XX
PS   Disclosure; Fig 1; 50pp; English.
XX
CC   The present invention describes a factor X analogue (I) which contains
CC   a modification between Glu226 and Ile235, relative to the 488 residue
CC   amino acid sequence given in AAB70411. (I) has haemostatic activity and
CC   can be used in gene therapy. (I) encoding polynucleotide (II) can be
CC   used to produce a drug, which is useful for treatment of patients with
CC   blood coagulation disorders, such as patients suffering from haemophilia,
CC   or haemophiliacs with inhibitory antibodies. Preparations containing a
CC   polypeptide with factor X/Xa activity are more readily activated by
CC   factor X1a or its derivative, which has high stability, without having
CC   to use one of the proteases used in prior art to activate the natural
CC   factor X, particularly one of animal origins, such as Russell's viper
CC   venom (RVV) or trypsin. The present sequence encodes human factor X,
CC   which is given in the exemplification of the present invention.
XX
SQ   Sequence 1467 BP; 363 A; 424 C; 444 G; 236 T; 0 other;

  Query Match             91.0%;  Score 91;  DB 22;  Length 1467;
  Best Local Similarity   94.9%;  Pred. No. 3e-22;
  Matches   94;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       2 GTCTGTCACAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 61
           |||  |  ||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db     737 GTCCCTGGCAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACTATTC 796

Qy      62 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 100
           |||||||||||||||||||||||||||||||||||||||
Db     797 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTC- 835


RESULT 14
AAA54031
ID   AAA54031 standard; DNA; 1507 BP.
XX
AC   AAA54031;
XX
DT   08-FEB-2001  (first entry)
XX
DE   Human factor X coding sequence.
XX
KW   Vitamin K dependent protein; VKDP; gamma-carboxylation; chimeric
KW   protein; fusion protein; coagulation factor; Factor X; Factor VII;
KW   Protein S; Factor IX; Protein C; prothrombin; blood clotting;
KW   haemophilia; human; ds
XX
OS   Homo sapiens.
XX
PN   WO200054787-A1.
XX
PD   21-SEP-2000.
XX
PF   16-MAR-2000; 2000WO-US06934.
XX
PR   16-MAR-1999;   99US-0124609.
XX
PA   (CHIL-) CHILDRENS HOSPITAL PHILADELPHIA.
PA   (UYNC-) UNIV NORTH CAROLINA.
XX
PI   High KA,  Camire RM,  Larson PJ,  Stafford DW;
XX
DR   WPI; 2000-638152/61.
XX
PT   Chimeric DNA for optimizing gamma carboxylation of vitamin K-dependent
PT   protein useful for treating diseases associated with the protein,
PT   comprises sequence encoding propeptide fused to sequence encoding the
PT   protein
XX
PS   Disclosure; Fig 6a; 60pp; English.
XX
CC   Efficient processing and release of mature two-chain factor X into
CC   the circulation requires: removal of the signal sequence; formation
CC   of disulfide bonds; modification of amino terminal glutamic acid
CC   residues, to gamma-carboxyglutamic acid; modification of one
CC   aspartic acid in the first epidermal growth factor (EGF) domain to
CC   Beta-hydroxyaspartic acid; addition of N- and O-linked
CC   oligosaccharides to the activation peptide; removal of an internal
CC   tripeptide to yield two-chain factor X and removal of the
CC   propeptide just prior to secretion. While some of these modifications
CC   do not appear essential for factor X function the removal of the
CC   signal sequence, propeptide, internal tripeptide and full
CC   gamma-carboxylation are all steps which are important requisites for
CC   the production of biologically active factor X/FXa. Isolated chimeric
CC   polynucleotides are described which encode a propeptide fused to a
CC   nucleic acid sequence encoding a vitamin K-dependent protein (VKDP).
CC   The fusion proteins encoded are vitamin K-dependent protein
CC   gamma-carboxylation enhancers and are useful for optimising the
CC   gamma-carboxylation of a VKDP to produce a fully gamma-carboxylated
CC   VKDP. The fusion proteins and recombinant cells expressing them are
CC   useful for alleviating a VKDP associated disease. The fusion
CC   constructs result in the production of fully gamma-carboxylated
CC   mature VKDPs, which are biologically active. The invention
CC   encompasses all combinations of propeptide sequences (modified or
CC   not) and VKDP's. This sequence encodes the signal, propeptide and
CC   mature protein sequence of human Factor X.
XX
SQ   Sequence 1507 BP; 394 A; 429 C; 446 G; 238 T; 0 other;

  Query Match             91.0%;  Score 91;  DB 21;  Length 1507;
  Best Local Similarity   94.9%;  Pred. No. 3e-22;
  Matches   94;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       2 GTCTGTCACAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 61
```

```
          |||  |  |||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db    737 GTCCCTGGCAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACTATTC 796

Qy     62 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 100
          |||||||||||||||||||||||||||||||||||||||
Db    797 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 835


RESULT 15
AAH57469
ID   AAH57469 standard; cDNA; 1887 BP.
XX
AC   AAH57469;
XX
DT   10-SEP-2001  (first entry)
XX
DE   Human liver cell specific cDNA sequence SEQ ID NO:309.
XX
KW   Human; tissue specific; diagnosis; brain; heart; skeletal muscle;
KW   lung; liver; uterus; ovary; stomach; intestine; kidney; pancreas; ss;
KW   metabolic disease; developmental disease; cytostatic; immunomodulatory;
KW   neuroprotective; gene therapy; cancer; immunopathology; neuropathology.
XX
OS   Homo sapiens.
XX
PN   WO200132927-A2.
XX
PD   10-MAY-2001.
XX
PF   02-NOV-2000; 2000WO-US30396.
XX
PR   04-NOV-1999;   99US-0163508.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Sornasse T,  Seilhamer JJ,  Watson GA;
XX
DR   WPI; 2001-291057/30.
XX
PT   New cell and tissue specific polynucleotides useful for diagnosis,
PT   prognosis or monitoring of treatments for disorders where the gene is
PT   associated with a cancer, immunopathology or neuropathology -
XX
PS   Claim 1; Page 233; 327pp; English.
XX
CC   AAH57161 to AAH57576 represent cell and tissue specific polynucleotide
CC   sequences (I). (I) can have cytostatic, immunomodulatory and
CC   neuroprotective activities, and can be used in gene therapy. (I) and
CC   proteins (II) encoded by then are used in high throughput screening
CC   assays to select DNA molecules, RNA molecules, peptide nucleic acids,
CC   mimetics, peptides, proteins, agonists, antagonists, antibodies or
CC   their fragments, immunoglobulins, inhibitors, drug compounds and
CC   pharmaceutical agents. Expression of (I) in a sample indicates the
CC   differentiation of embryonic stem cells into a tissue selected from
CC   brain, heart, kidney, liver, lung, skeletal muscle or pancreatic
CC   tissues. (I) and (II) are used to produce an expression profile that
CC   defines a metabolic or developmental process, treatment, condition,
CC   disease or disorder. The gene profile can be used for diagnosis,
CC   prognosis or monitoring of treatments and for investigating a
CC   predisposition to a disorder where the gene is associated with a
CC   cancer, immunopathology or neuropathology.
XX
SQ   Sequence 1887 BP; 467 A; 549 C; 544 G; 327 T; 0 other;

  Query Match             91.0%;  Score 91;  DB 22;  Length 1887;
  Best Local Similarity   94.9%;  Pred. No. 3.3e-22;
  Matches   94;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      2 GTCTGTCACAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 61
          |||  |  |||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db    792 GTCCCTGGCAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACTATTC 851

Qy     62 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 100
          ||||||||||||||||||||||||||||||||||||||||
Db    852 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 890


Search completed: January 15, 2003, 19:50:17
Job time : 153 secs
```

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: January 15, 2003, 17:58:10 ; Search time 31 Seconds
(without alignments)
989.280 Million cell updates/sec

Title: L00396_COPY_1_100
Perfect score: 100
Sequence: 1 CGTCTGTCACAGGCCCTGCT.....AACGGCAGCCCACTGTCTCT 100

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	92.6	92.6	725	3	US-09-197-801-12	Sequence 12, Appl
2	92.6	92.6	725	4	US-09-551-028-12	Sequence 12, Appl
3	92.6	92.6	1126	2	US-08-479-733A-27	Sequence 27, Appl
4	92.6	92.6	1126	3	US-08-487-427-27	Sequence 27, Appl
5	92.6	92.6	1126	3	US-08-479-727A-27	Sequence 27, Appl
6	92.6	92.6	1126	3	US-08-482-369A-27	Sequence 27, Appl
7	92.6	92.6	1126	5	PCT-US95-07439-27	Sequence 27, Appl
8	92.6	92.6	1404	4	US-09-202-101-15	Sequence 15, Appl
9	91	91.0	1500	1	US-08-487-037-4	Sequence 4, Appli
c 10	75.2	75.2	591	4	US-09-280-116-138	Sequence 138, App
11	63.4	63.4	1554	1	US-08-469-486-1	Sequence 1, Appli
12	63.4	63.4	1554	2	US-08-469-658-1	Sequence 1, Appli
13	34.2	34.2	2038	4	US-09-008-271A-18	Sequence 18, Appl
14	34.2	34.2	2079	4	US-09-656-002-1	Sequence 1, Appli
15	32	32.0	864	3	US-08-906-769-138	Sequence 138, App
16	32	32.0	864	3	US-08-906-616-138	Sequence 138, App
17	32	32.0	864	3	US-08-639-075A-138	Sequence 138, App
18	32	32.0	864	4	US-09-012-431-138	Sequence 138, App
19	32	32.0	864	4	US-09-012-692-138	Sequence 138, App
20	32	32.0	864	4	US-08-906-613-138	Sequence 138, App
21	30.8	30.8	855	3	US-08-906-769-130	Sequence 130, App
22	30.8	30.8	855	3	US-08-906-616-130	Sequence 130, App
23	30.8	30.8	855	3	US-08-639-075A-130	Sequence 130, App
24	30.8	30.8	855	4	US-09-012-431-130	Sequence 130, App
25	30.8	30.8	855	4	US-09-012-692-130	Sequence 130, App
26	30.8	30.8	855	4	US-08-906-613-130	Sequence 130, App
27	29.4	29.4	758	3	US-08-906-769-126	Sequence 126, App
28	29.4	29.4	758	3	US-08-906-616-126	Sequence 126, App
29	29.4	29.4	758	3	US-08-639-075A-126	Sequence 126, App
30	29.4	29.4	758	4	US-09-012-431-126	Sequence 126, App
31	29.4	29.4	758	4	US-09-012-692-126	Sequence 126, App
32	29.4	29.4	758	4	US-08-906-613-126	Sequence 126, App
33	28.8	28.8	258	1	US-08-485-455D-52	Sequence 52, Appl
34	28.8	28.8	258	2	US-08-482-130C-52	Sequence 52, Appl
35	28.8	28.8	258	2	US-08-484-211C-52	Sequence 52, Appl
36	28.8	28.8	258	3	US-08-906-769-52	Sequence 52, Appl
37	28.8	28.8	258	3	US-08-906-616-52	Sequence 52, Appl
38	28.8	28.8	258	3	US-08-817-795-52	Sequence 52, Appl
39	28.8	28.8	258	3	US-08-485-443B-52	Sequence 52, Appl
40	28.8	28.8	258	3	US-08-639-075A-52	Sequence 52, Appl
41	28.8	28.8	258	4	US-09-012-431-52	Sequence 52, Appl
42	28.8	28.8	258	4	US-09-012-692-52	Sequence 52, Appl
43	28.8	28.8	258	4	US-08-906-613-52	Sequence 52, Appl
44	28.8	28.8	258	5	PCT-US95-14442A-52	Sequence 52, Appl
45	28	28.0	1389	4	US-09-202-101-16	Sequence 16, Appl

ALIGNMENTS

RESULT 1
US-09-197-801-12
; Sequence 12, Application US/09197801B
; Patent No. 6159722
; GENERAL INFORMATION:
; APPLICANT: Kopetzki, Erhard
; APPLICANT: Hopfner, Karl-Peter
; APPLICANT: Engh, Richard
; APPLICANT: Bode, Wolfram
; APPLICANT: Huber, Robert
; TITLE OF INVENTION: Chimeric Serine Proteases
; FILE REFERENCE: 20119
; CURRENT APPLICATION NUMBER: US/09/197,801B
; CURRENT FILING DATE: 1998-11-23
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 725
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-197-801-12

Query Match 92.6%; Score 92.6; DB 3; Length 725;
Best Local Similarity 96.0%; Pred. No. 4e-25;
Matches 95; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy       2 GTCTGTCACAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 61
           ||| | ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      65 GTCCCTGGCAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 124

Qy      62 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 100
           ||||||||||||||||||||||||||||||||||||||||
Db     125 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 163
```

RESULT 2
US-09-551-028-12
; Sequence 12, Application US/09551028
; Patent No. 6171842
; GENERAL INFORMATION:
; APPLICANT: Kopetzki, Erhard
; APPLICANT: Hopfner, Karl-Peter
; APPLICANT: Engh, Richard
; APPLICANT: Bode, Wolfram
; APPLICANT: Huber, Robert
; TITLE OF INVENTION: Chimeric Serine Proteases
; FILE REFERENCE: 20119
; CURRENT APPLICATION NUMBER: US/09/551,028
; CURRENT FILING DATE: 2000-04-17
; PRIOR APPLICATION NUMBER: US/09/197,801

```
;  PRIOR FILING DATE:   1998-11-23
;  NUMBER OF SEQ ID NOS:   15
;  SOFTWARE:   PatentIn Ver. 2.0
; SEQ ID NO 12
;   LENGTH: 725
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-551-028-12

  Query Match             92.6%;  Score 92.6;  DB 4;  Length 725;
  Best Local Similarity   96.0%;  Pred. No. 4e-25;
  Matches   95;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 GTCTGTCACAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 61
           |||  |  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      65 GTCCCTGGCAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 124

Qy      62 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 100
           ||||||||||||||||||||||||||||||||||||||||
Db     125 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 163


RESULT 3
US-08-479-733A-27
; Sequence 27, Application US/08479733A
; Patent No. 5877289
;  GENERAL INFORMATION:
;    APPLICANT:  Thorpe, Philip E.
;    APPLICANT:  Edgington, Thomas S.
;    TITLE OF INVENTION:  Methods and Compositions for the
;    TITLE OF INVENTION:  Specific Coagulation of Vasculature
;    NUMBER OF SEQUENCES:  32
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  US
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/479,733A
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/273,567
;      FILING DATE:  11-JUL-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Parker, David L.
;      REGISTRATION NUMBER:  32,165
;      REFERENCE/DOCKET NUMBER:  UTSD:459/PAR
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  512/418-3000
;      TELEFAX:  512/474-7577
;      TELEX:  N/A
;  INFORMATION FOR SEQ ID NO:  27:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1126 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-479-733A-27

  Query Match             92.6%;  Score 92.6;  DB 2;  Length 1126;
  Best Local Similarity   96.0%;  Pred. No. 4.7e-25;
  Matches   95;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 GTCTGTCACAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 61
           |||  |  ||||||||||||||||||||||||||||||||||||||||||||||||||||

Db     395 GTCCCTGGCAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 454

Qy      62 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 100
           ||||||||||||||||||||||||||||||||||||||||
Db     455 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 493


RESULT 4
US-08-487-427-27
; Sequence 27, Application US/08487427
; Patent No. 6004555
;  GENERAL INFORMATION:
;    APPLICANT:  Thorpe, Philip E.
;    APPLICANT:  Edgington, Thomas S.
;    TITLE OF INVENTION:  Methods and Compositions for the
;    TITLE OF INVENTION:  Specific Coagulation of Vasculature
;    NUMBER OF SEQUENCES:  27
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  US
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/487,427
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/273,567
;      FILING DATE:  11-JUN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Parker, David L.
;      REGISTRATION NUMBER:  32,165
;      REFERENCE/DOCKET NUMBER:  UTSD:457/PAR
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  512/418-3000
;      TELEFAX:  512/474-7577
;      TELEX:  N/A
;  INFORMATION FOR SEQ ID NO:  27:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1126 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-487-427-27

  Query Match             92.6%;  Score 92.6;  DB 3;  Length 1126;
  Best Local Similarity   96.0%;  Pred. No. 4.7e-25;
  Matches   95;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 GTCTGTCACAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 61
           |||  |  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     395 GTCCCTGGCAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 454

Qy      62 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 100
           ||||||||||||||||||||||||||||||||||||||||
Db     455 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 493


RESULT 5
US-08-479-727A-27
; Sequence 27, Application US/08479727A
; Patent No. 6036955
;  GENERAL INFORMATION:
;    APPLICANT:  Thorpe, Philip E.
;    TITLE OF INVENTION:  Methods and Compositions for the
;    TITLE OF INVENTION:  Specific Coagulation of Vasculature
```

```
;    NUMBER OF SEQUENCES:  32
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  US
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/479,727A
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/273,567
;      FILING DATE:  11-JUL-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Parker, David L.
;      REGISTRATION NUMBER:  32,165
;      REFERENCE/DOCKET NUMBER:  UTSD:461/PAR
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  512/418-3000
;      TELEFAX:  512/474-7577
;      TELEX:  N/A
;  INFORMATION FOR SEQ ID NO:  27:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1126 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-479-727A-27

  Query Match             92.6%;  Score 92.6;  DB 3;  Length 1126;
  Best Local Similarity   96.0%;  Pred. No. 4.7e-25;
  Matches   95;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 GTCTGTCACAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 61
           |||  |  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     395 GTCCCTGGCAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 454

Qy      62 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 100
           |||||||||||||||||||||||||||||||||||||||
Db     455 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 493


RESULT 6
US-08-482-369A-27
; Sequence 27, Application US/08482369A
; Patent No. 6093399
;  GENERAL INFORMATION:
;    APPLICANT:  Thorpe, Philip E.
;    APPLICANT:  Edgington, Thomas S.
;    TITLE OF INVENTION:  Methods and Compositions for the Specific
;    TITLE OF INVENTION:  Coagulation of Vasculature
;    NUMBER OF SEQUENCES:  32
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  US
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/482,369A
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/273,567
;      FILING DATE:  11-JUL-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Parker, David L.
;      REGISTRATION NUMBER:  32,165
;      REFERENCE/DOCKET NUMBER:  UTSD:433/PAR
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  512/418-3000
;      TELEFAX:  512/474-7577
;      TELEX:  N/A
;  INFORMATION FOR SEQ ID NO:  27:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1126 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-482-369A-27

  Query Match             92.6%;  Score 92.6;  DB 3;  Length 1126;
  Best Local Similarity   96.0%;  Pred. No. 4.7e-25;
  Matches   95;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 GTCTGTCACAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 61
           |||  |  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     395 GTCCCTGGCAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 454

Qy      62 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 100
           |||||||||||||||||||||||||||||||||||||||
Db     455 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 493


RESULT 7
PCT-US95-07439-27
; Sequence 27, Application PC/TUS9507439
;  GENERAL INFORMATION:
;    APPLICANT:
;    APPLICANT:  NAME:  BOARD OF REGENTS, THE UNIVERSITY OF
;    APPLICANT:  TEXAS SYSTEM
;    APPLICANT:  STREET:  201 West 7th Street
;    APPLICANT:  CITY:  Austin
;    APPLICANT:  STATE:  Texas
;    APPLICANT:  COUNTRY:  United States of America
;    APPLICANT:  POSTAL CODE:  78701
;    APPLICANT:  TELEPHONE NO: (512)499-4462
;    APPLICANT:  TELEFAX:  (512)499-4523
;    APPLICANT:  NAME:  THE SCRIPPS RESEARCH INSTITUTE
;    APPLICANT:  STREET:  10666 North Torrey Pines Road
;    APPLICANT:  CITY:  LaJolla
;    APPLICANT:  STATE:  California
;    APPLICANT:  COUNTRY:  United States of America
;    APPLICANT:  POSTAL CODE:  92037
;    TITLE OF INVENTION:  METHODS AND COMPOSITIONS
;    TITLE OF INVENTION:  FOR THE SPECIFIC
;    TITLE OF INVENTION:  COAGULATION OF VASCULATURE
;    NUMBER OF SEQUENCES:  27
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P. O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  USA
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS, ASCII
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/07439
;      FILING DATE:  Concurrently herewith
;      CLASSIFICATION:
```

```
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/273,567
;      FILING DATE:  11-JUN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  PARKER, DAVID L.
;      REGISTRATION NUMBER:  32,165
;      REFERENCE/DOCKET NUMBER:  UTFD433P--
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (512) 418-3000
;      TELEFAX:  (713) 789-2679
;      TELEX:  79-0924
;  INFORMATION FOR SEQ ID NO:  27:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1126 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
PCT-US95-07439-27

  Query Match             92.6%;  Score 92.6;  DB 5;  Length 1126;
  Best Local Similarity   96.0%;  Pred. No. 4.7e-25;
  Matches   95;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 GTCTGTCACAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 61
           |||  |  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     395 GTCCCTGGCAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 454

Qy      62 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 100
           |||||||||||||||||||||||||||||||||||||||
Db     455 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 493


RESULT 8
US-09-202-101-15
; Sequence 15, Application US/09202101
; Patent No. 6277618
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  Recombinant blood-coagulation proteases
;    NUMBER OF SEQUENCES:  17
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30B (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/202,101
;      FILING DATE:
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1404 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-202-101-15

  Query Match             92.6%;  Score 92.6;  DB 4;  Length 1404;
  Best Local Similarity   96.0%;  Pred. No. 5.1e-25;
  Matches   95;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 GTCTGTCACAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 61
           |||  |  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     683 GTCCCTGGCAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 742

Qy      62 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 100
           |||||||||||||||||||||||||||||||||||||||
Db     743 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 781


RESULT 9
US-08-487-037-4
; Sequence 4, Application US/08487037
; Patent No. 5795863
;  GENERAL INFORMATION:
;    APPLICANT:  Wolf, David L.
;    TITLE OF INVENTION:  RECOMBINANT AGENTS AFFECTING THROMBOSIS
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  2000 Pennsylvania Avenue, NW
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20006-1812
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/487,037
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Adler, Reid G.
;      REGISTRATION NUMBER:  30,988
;      REFERENCE/DOCKET NUMBER:  2803-0002.02
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 887-1500
;      TELEFAX:  (202) 887-0763
;      TELEX:  90-4030
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1500 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-487-037-4

  Query Match             ..0%;  Score 91;  DB 1;  Length 1500;
  Best Local Similarity   94.9%;  Pred. No. 2e-24;
  Matches   94;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       2 GTCTGTCACAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTC 61
           |||  |  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     769 GTCCCTGGCAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACTATTC 828

Qy      62 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 100
           |||||||||||||||||||||||||||||||||||||||
Db     829 TGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCT 867


RESULT 10
US-09-280-116-138/c
; Sequence 138, Application US/09280116A
; Patent No. 6331427
; GENERAL INFORMATION:
;  APPLICANT: Robison, Keith E.
;  TITLE OF INVENTION: Nucleic Acid Molecules Encoding Human Protease Homologs
;  FILE REFERENCE: 5800-24, 035800/176965
;  CURRENT APPLICATION NUMBER: US/09/280,116A
;  CURRENT FILING DATE: 1999-03-26
;  NUMBER OF SEQ ID NOS: 268
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 138
;   LENGTH: 591
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: trypsin-like serine proteases
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(591)
;   OTHER INFORMATION: n = a, t, c or g
US-09-280-116-138
```

```
  Query Match           75.2%;  Score 75.2;  DB 4;  Length 591;
  Best Local Similarity  95.7%;  Pred. No. 1.1e-18;
  Matches   88;  Conservative    0;  Mismatches    3;  Indels    1;  Gaps    1;

Qy       6 GTCACAGGCCCTGCTCATCAATGAGGAAAACG-AGGGTTTCTGTGGTGGAACCATTCTGA 64
           |||||||||||||||||||||||||||||||| |||||||||||||||||||| |||||||
Db     499 GTCACAGGCCCTGCTCATCAATGAGGAAAACGAAGGGTTTCTGTGGTGGAACTATTCTGA 440

Qy      65 GCGAGTTCTACATCCTAACGGCAGCCCACTGT 96
           ||||||||||||||  ||||||||||||||||
Db     439 GCGAGTTCTACATCCCTACGGCAGCCCACTGT 408


RESULT 11
US-08-469-486-1
; Sequence 1, Application US/08469486
; Patent No. 5739281
;  GENERAL INFORMATION:
;    APPLICANT:  Thoegersen, Hans Christian
;    APPLICANT:  Holtet, Thor Las
;    APPLICANT:  Etzerodt, Michael
;    TITLE OF INVENTION:  Improved method for the refolding of
;    TITLE OF INVENTION:  proteins
;    NUMBER OF SEQUENCES:  58
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version
;      SOFTWARE:  #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/469,486
;      FILING DATE:
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/192,060
;      FILING DATE:  February 4, 1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Paul T. Clark
;      REGISTRATION NUMBER:  30,162
;      REFERENCE/DOCKET NUMBER:  06363/002001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617 542 5070
;      TELEFAX:  617 542 8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1554 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  YES
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Bos taurus
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  76..1551
US-08-469-486-1

  Query Match           63.4%;  Score 63.4;  DB 1;  Length 1554;
  Best Local Similarity  82.0%;  Pred. No. 3.6e-14;
  Matches   73;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy      10 CAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTCTGAGCGAG 69
           ||||| |||||  |||| || || |||||||| |||| || || ||||| |||| ||||
Db     817 CAGGCTCTGCTGGTCAACGAAGAGAACGAGGGATTCTGCGGGGGCACCATCCTGAACGAG 876

Qy      70 TTCTACATCCTAACGGCAGCCCACTGTCT 98
           |||||| |||| ||||| |||||||| ||
Db     877 TTCTACGTCCTCACGGCTGCCCACTGCCT 905


RESULT 12
US-08-469-658-1
; Sequence 1, Application US/08469658
; Patent No. 5917018
;  GENERAL INFORMATION:
;    APPLICANT:  Thoegersen, Hans Christian
;    APPLICANT:  Holtet, Thor Las
;    APPLICANT:  Etzerodt, Michael
;    TITLE OF INVENTION:  IMPROVED METHOD FOR THE REFOLDING OF
;    TITLE OF INVENTION:  PROTEINS
;    NUMBER OF SEQUENCES:  58
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version
;      SOFTWARE:  #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/469,658
;      FILING DATE:  June 5, 1995
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/192,060
;      FILING DATE:  February 4, 1994
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Paul T. Clark
;      REGISTRATION NUMBER:  30,162
;      REFERENCE/DOCKET NUMBER:  06363/002002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617 542 5070
;      TELEFAX:  617 542 8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1554 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  YES
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Bos taurus
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  76..1551
US-08-469-658-1

  Query Match           63.4%;  Score 63.4;  DB 2;  Length 1554;
  Best Local Similarity  82.0%;  Pred. No. 3.6e-14;
  Matches   73;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy      10 CAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTCTGAGCGAG 69
           ||||| |||||  |||| || || |||||||| |||| || || ||||| |||| ||||
Db     817 CAGGCTCTGCTGGTCAACGAAGAGAACGAGGGATTCTGCGGGGGCACCATCCTGAACGAG 876
```

```
Qy      70 TTCTACATCCTAACGGCAGCCCACTGTCT 98
           |||||| |||| ||||| |||||||| ||
Db     877 TTCTACGTCCTCACGGCTGCCCACTGCCT 905


RESULT 13
US-09-008-271A-18
; Sequence 18, Application US/09008271A
; Patent No. 6203979
;  GENERAL INFORMATION:
;       APPLICANT: Bandman, Olga
;                  Hillman, Jennifer L.
;                  Yue, Henry
;                  Guegler, Karl J.
;                  Corley, Neil C.
;                  Tang, Tom Y.
;                  Shah, Purvi
;       TITLE OF INVENTION: HUMAN PROTEASE MOLECULES
;       NUMBER OF SEQUENCES: 24
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Incyte Pharmaceuticals, Inc.
;            STREET: 3174 Porter Dr.
;            CITY: Palo Alto
;            STATE: CA
;            COUNTRY: USA
;            ZIP: 94304
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Diskette
;            COMPUTER: IBM Compatible
;            OPERATING SYSTEM: DOS
;            SOFTWARE: FastSEQ for Windows Version 2.0
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/008,271A
;            FILING DATE: 16-Jan-1998
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: <Unknown>
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Mohan-Peterson, Sheela
;            REGISTRATION NUMBER: 41,201
;            REFERENCE/DOCKET NUMBER: PF-0458 US
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 650-855-0555
;            TELEFAX: 650-845-4166
;  INFORMATION FOR SEQ ID NO: 18:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 2038 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       IMMEDIATE SOURCE:
;            LIBRARY: COLNNOT13
;            CLONE: 1337018
;       SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-09-008-271A-18

  Query Match             34.2%;  Score 34.2;  DB 4;  Length 2038;
  Best Local Similarity   64.6%;  Pred. No. 0.0026;
  Matches   51;  Conservative    0;  Mismatches   28;  Indels    0;  Gaps    0;

Qy      21 CATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCAT CTGAGCGAGTT CTACATCCT 80
           |||| |  | || ||  ||    ||||||| || | ||| |||  |     ||   ||||
Db     856 CATCCAGTACGACAAACAGCACGTCTGTGGAGGGAGCATCCTGGACCCCCACTGGGTCCT 915

Qy      81 AACGGCAGCCCACTGTCTC 99
           |||||||||||||||  ||
Db     916 CACGGCAGCCCACTGCTTC 934


RESULT 14
US-09-656-002-1
; Sequence 1, Application US/09656002
; Patent No. 6455668
;  GENERAL INFORMATION:
;   APPLICANT: Mack, David
;   APPLICANT:  Gish, Kurt
;   APPLICANT:  Wilson, Keith
;   TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSING COLORECTAL CANCER, COMPOSITIONS, A
;   TITLE OF INVENTION:  OF SCREENING FOR COLORECTAL CANCER MODULATORS
;   FILE REFERENCE: A-69108/DJB/JJD/AMS
;   CURRENT APPLICATION NUMBER: US/09/656,002
;   CURRENT FILING DATE:  2000-09-06
;   PRIOR APPLICATION NUMBER: US 09/525,993
;   PRIOR FILING DATE: 2000-03-15
;   PRIOR APPLICATION NUMBER: US 09/493,444
;   PRIOR FILING DATE: 2000-01-28
;   PRIOR APPLICATION NUMBER: PCT/US 00/07044
;   PRIOR FILING DATE: 2000-03-15
;   NUMBER OF SEQ ID NOS: 3
;   SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
;   LENGTH: 2079
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-656-002-1

  Query Match             34.2%;  Score 34.2;  DB 4;  Length 2079;
  Best Local Similarity   64.6%;  Pred. No. 0.0027;
  Matches   51;  Conservative    0;  Mismatches   28;  Indels    0;  Gaps    0;

Qy      21 CATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTCTGAGCGAGTTCTACATCCT 80
           |||| |  | || ||  ||    ||||||| || | ||| |||  |     ||   ||||
Db     871 CATCCAGTACGACAAACAGCACGTCTGTGGAGGGAGCATCCTGGACCCCCACTGGGTCCT 930

Qy      81 AACGGCAGCCCACTGTCTC 99
           |||||||||||||||  ||
Db     931 CACGGCAGCCCACTGCTTC 949


RESULT 15
US-08-906-769-138
; Sequence 138, Application US/08906769
; Patent No. 6077687
;  GENERAL INFORMATION:
;    APPLICANT:  Grieve, Robert B.
;    APPLICANT:  Rushlow, Keith E.
;    APPLICANT:  Wu Hunter, Shirley
;    APPLICANT:  Frank, Glenn R.
;    APPLICANT:  Stiegler, Gary
;    APPLICANT:  Gaines, Patrick J.
;    APPLICANT:  Silver, Gary
;    TITLE OF INVENTION:  FLEA PROTEASE PROTEINS, NUCLEIC ACID
;    TITLE OF INVENTION:  MOLECULES AND USES THEREOF
;    NUMBER OF SEQUENCES:  190
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sheridan Ross & McIntosh
;      STREET:  1700 Lincoln Street, Suite 3500
;      CITY:  Denver
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80203
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/906,769
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/639,075
;      FILING DATE:  24-APR-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Connell, Gary J.
;      REGISTRATION NUMBER:  32,020
```

```
;      REFERENCE/DOCKET NUMBER:  2618-25-C2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (303) 863-9700
;      TELEFAX:  (303) 863-0223
;  INFORMATION FOR SEQ ID NO:  138:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  864 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  2..781
;      OTHER INFORMATION:  /note= "At pos. bp 456, change G to
;      OTHER INFORMATION:  K; at pos. bp 504, change A to R.  At pos. aa 152 and 168,
;      OTHER INFORMATION:  substitute Xaa."
US-08-906-769-138

  Query Match            32.0%;  Score 32;  DB 3;  Length 864;
  Best Local Similarity  58.3%;  Pred. No. 0.013;
  Matches   56;  Conservative    0;  Mismatches   40;  Indels    0;  Gaps    0;

Qy       3 TCTGTCACAGGCCCTGCTCATCAATGAGGAAAACGAGGGTTTCTGTGGTGGAACCATTCT 62
           || ||  || |     || |  |||  |||    |||  ||||||||||||  ||||  |
Db     145 TCCGTATCAAGTTTCACTGAGAAATCGGGACCTAGAGCATTTCTGTGGTGGCTCCATCTT 204

Qy      63 GAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCT 98
           | | |    |  ||  | ||||| || || || ||
Db     205 AAACAAACGTTGGATTGTGACGGCTGCACATTGCCT 240


Search completed: January 15, 2003, 20:51:15
Job time : 33 secs
```